

STIC-Biotech/ChemLib

168297

From: Chan, Christina
Sent: Tuesday, October 11, 2005 1:40 PM
To: Dunston, Jennifer; STIC-Biotech/ChemLib
Subject: RE: Sequence Search 10/826523

Please rush. Thanks Chris

Chris Chan
TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841
Remsen, 3E89

RECEIVED
OCT 11 2005
STIC/CHEN, Jennifer
(STIC)

-----Original Message-----

From: Dunston, Jennifer
Sent: Tuesday, October 11, 2005 1:19 PM
To: Chan, Christina
Subject: Sequence Search 10/826523

Please RUSH this search. It is for an after final amendment.

Please do a sequence search for the nucleic acid sequence of SEQ ID NO: 40 against the commercial and interference nucleotide databases.

The length of SEQ ID NO: 40 is 707 nucleotides.

Jennifer Dunston, Ph.D.
USPTO Art Unit 1636
REM 2B76
Mailbox: REM 2C70
(571) 272-2916

Searcher: _____
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: _____
Searcher Prep Time: _____
Online Time: _____

Type of Search
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other (Specify): _____

100

Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 12, 2005, 22:41:13 ; Search time 5848 Seconds
(without alignments)
847.469 Million cell updates/sec

Title: US-10-826-523-40

Perfect score: 707

Sequence: 1 ggatcccatcgctcaatttt.....aaaattgacgcatgggatcc 707

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 8765912 seqs, 3504951483 residues

Total number of hits satisfying chosen parameters: 17531824

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description
1	707	100.0	707	US-10-001-189-40 Sequence 40, Appl
2	707	100.0	4613	US-10-001-189-46 Sequence 46, Appl
3	707	100.0	8999	US-10-001-189-48 Sequence 48, Appl
4	707	100.0	9012	US-10-001-189-49 Sequence 49, Appl
5	707	100.0	9013	US-10-001-189-50 Sequence 50, Appl
6	685	96.9	4943	US-10-001-189-54 Sequence 54, Appl
7	685	96.9	4944	US-10-001-189-55 Sequence 55, Appl

8	685	96.9	4944	13	US-10-001-189-56	Sequence 56, Appl
9	580.6	82.1	4164	22	US-10-220-335-158	Sequence 158, Appl
c 10	566.2	80.1	26565	16	US-10-085-959-91	Sequence 91, Appl
c 11	566.2	80.1	46819	14	US-10-114-170-72	Sequence 72, Appl
c 12	561.4	79.4	22306	14	US-10-114-170-251	Sequence 251, Appl
c 13	519.2	73.4	3662	13	US-10-001-189-41	Sequence 41, Appl
c 14	511	72.3	4941	13	US-10-001-189-53	Sequence 53, Appl
c 15	511	72.3	4951	13	US-10-001-189-51	Sequence 51, Appl
c 16	511	72.3	4952	13	US-10-001-189-52	Sequence 52, Appl
c 17	425.8	60.2	1427	20	US-10-363-345A-37517	Sequence 37517, A
c 18	425.8	60.2	1427	20	US-10-363-345A-37518	Sequence 37518, A
c 19	425.8	60.2	1427	21	US-10-363-483A-37517	Sequence 37517, A
c 20	425.8	60.2	1427	21	US-10-363-483A-37518	Sequence 37518, A
c 21	392.6	55.5	3822	22	US-10-450-763-30289	Sequence 30289, A
c 22	382.6	54.1	1427	20	US-10-363-345A-37519	Sequence 37519, A
c 23	382.6	54.1	1427	21	US-10-363-483A-37520	Sequence 37520, A
c 24	382.6	54.1	1427	21	US-10-363-483A-37519	Sequence 37519, A
c 25	382.6	54.1	1427	21	US-10-363-483A-37520	Sequence 37520, A
c 26	223	31.5	34063	14	US-10-114-170-96	Sequence 96, Appl
c 27	222.2	31.4	72480	17	US-10-418-831-2	Sequence 2, Appl
c 28	171.2	24.2	822	22	US-10-450-763-15978	Sequence 15978, A
c 29	96	13.6	847	10	US-09-798-889-27	Sequence 27, Appl
c 30	96	13.6	847	18	US-10-633-680-27	Sequence 27, Appl
c 31	78	11.0	9984	13	US-10-001-189-63	Sequence 63, Appl
c 32	72.8	10.3	7670	13	US-10-001-189-57	Sequence 57, Appl
c 33	64.6	9.1	1275	22	US-10-450-763-23043	Sequence 23043, A
c 34	60.8	5.8	1523	22	US-10-450-763-15121	Sequence 15121, A
c 35	40	5.7	43	13	US-10-001-189-19	Sequence 19, Appl
c 36	40	5.7	3089	20	US-10-425-115-49726	Sequence 49726, A
c 37	39.8	5.6	676	20	US-10-425-115-91844	Sequence 91844, A
c 38	39.8	5.6	1673	18	US-10-425-114-21174	Sequence 21174, A
c 39	39.8	5.6	1796	18	US-10-425-114-29053	Sequence 29053, A
c 40	39.8	5.6	1872	20	US-10-425-115-91842	Sequence 91842, A
c 41	39	5.5	39	13	US-10-001-189-16	Sequence 16, Appl
c 42	38.4	5.4	1473	17	US-10-282-122A-20134	Sequence 20134, A
c 43	36.8	5.2	4187	19	US-10-437-963-91699	Sequence 91699, A
c 44	36.2	5.1	549	21	US-10-660-811A-33	Sequence 23, Appl
c 45	36.2	5.1	1403	19	US-10-767-701-13724	Sequence 13724, A

ALIGNMENTS

RESULT 1

US-10-001-189-40
; Sequence 40, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSCRIPTION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; TITLE OF INVENTION: VECTOR PIGGYBAC
; FILE REFERENCE: 835910-32098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 707
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: ITR Cartridge
; OTHER INFORMATION: sequence
US-10-001-189-40

Query Match 100.0%; Score 707; DB 13; Length 707;

```
Best Local Similarity 100.0%; Pred. No. 2.1e-235;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATCCCATGGTCAATTTTACGAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
Db 1 GGATCCCATGGTCAATTTTACGAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60

Qy 61 ATCATATCGTGGGTCTTTTTCGGGCTCAGTCACGCCCAAGCTGGCGCTATCTGGGCA 120
Db 61 ATCATATCGTGGGTCTTTTTCGGGCTCAGTCACGCCCAAGCTGGCGCTATCTGGGCA 120

Qy 121 TCGGGAGGAGAAAGCCGCTGCTTTTCCCGAGGTTGAAGCGGCATGGAAGAGTTTG 180
Db 121 TCGGGAGGAGAAAGCCGCTGCTTTTCCCGAGGTTGAAGCGGCATGGAAGAGTTTG 180

Qy 181 CCGAGGATGACTGCTGCTGCAATGAGTTGAGCGGAAACCGACGTTTACCATGATGATTC 240
Db 181 CCGAGGATGACTGCTGCTGCAATGAGTTGAGCGGAAACCGACGTTTACCATGATGATTC 240

Qy 241 GGGAGGTGTGGCCATGACACGCTTTTAAACGGTGAACCTGTTTCAGGCCACCTGGGATA 300
Db 241 GGGAGGTGTGGCCATGACACGCTTTTAAACGGTGAACCTGTTTCAGGCCACCTGGGATA 300

Qy 301 CCAGTTGCTCGCGGCTTTTCCGGACACAGTTCCGGATGGTTCAGCCGAGCGCATCAGCA 360
Db 301 CCAGTTGCTCGCGGCTTTTCCGGACACAGTTCCGGATGGTTCAGCCGAGCGCATCAGCA 360

Qy 361 ACCGGAACAATACCGGACAGCGGAACTGCGCTGCGGTTGAGATTAATGACAGCG 420
Db 361 ACCGGAACAATACCGGACAGCGGAACTGCGCTGCGGTTGAGATTAATGACAGCG 420

Qy 421 GTGGCGGCTGGGATATTACGTACGAGGAGCGGTATCCTGGCTGGATGCCGAGAAAT 480
Db 421 GTGGCGGCTGGGATATTACGTACGAGGAGCGGTATCCTGGCTGGATGCCGAGAAAT 480

Qy 481 GGACATGATACCCCGTGAGTTTACCCGGCGGCGGCTCGTTCAATTCACGTTTTTGAAC 540
Db 481 GGACATGATACCCCGTGAGTTTACCCGGCGGCGGCTCGTTCAATTCACGTTTTTGAAC 540

Qy 541 CCGTGAGGAGCGGAGACTGCGGTCGAAATGTGTTTACAGCTGATGAGCGAGTGA 600
Db 541 CCGTGAGGAGCGGAGACTGCGGTCGAAATGTGTTTACAGCTGATGAGCGAGTGA 600

Qy 601 AGATGCTCGACACGCTGACAGAACCGCAGCTAGATTAAACCTAGAAAGATTAATCATTTG 660
Db 601 AGATGCTCGACACGCTGACAGAACCGCAGCTAGATTAAACCTAGAAAGATTAATCATTTG 660

Qy 661 TGACGTACGTTAAAGATAATCATCGTAAAAATTGACGCATGGGATCC 707
Db 661 TGACGTACGTTAAAGATAATCATCGTAAAAATTGACGCATGGGATCC 707
```

```
RESULT 2
US-10-001-189-46
; Sequence 46, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
```

```
; LENGTH: 4613
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: pCRII-ITR
; OTHER INFORMATION: sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (34)..(922)
US-10-001-189-46

Query Match 100.0%; Score 707; DB 13; Length 4613;
Best Local Similarity 100.0%; Pred. No. 5.6e-235;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGATCCCATGGTCAATTTTACGAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
Db 294 GGATCCCATGGTCAATTTTACGAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 353

Qy 61 ATCATATCGTGGGTCTTTTTCGGGCTCAGTCACGCCCAAGCTGGCGCTATCTGGGCA 120
Db 354 ATCATATCGTGGGTCTTTTTCGGGCTCAGTCACGCCCAAGCTGGCGCTATCTGGGCA 413

Qy 121 TCGGGAGGAGAAAGCCGCTGCTTTTCCCGAGGTTGAAGCGGCATGGAAGAGTTTG 180
Db 414 TCGGGAGGAGAAAGCCGCTGCTTTTCCCGAGGTTGAAGCGGCATGGAAGAGTTTG 473

Qy 181 CCGAGGATGACTGCTGCTGCAATGAGTTGAGCGGAAACCGACGTTTACCATGATGATTC 240
Db 474 CCGAGGATGACTGCTGCTGCAATGAGTTGAGCGGAAACCGACGTTTACCATGATGATTC 533

Qy 241 GGGAGGTGTGGCCATGACACGCTTTTAAACGGTGAACCTGTTTCAGGCCACCTGGGATA 300
Db 534 GGGAGGTGTGGCCATGACACGCTTTTAAACGGTGAACCTGTTTCAGGCCACCTGGGATA 593

Qy 301 CCAGTTGCTCGCGGCTTTTCCGGACACAGTTCCGGATGGTTCAGCCGAGCGCATCAGCA 360
Db 594 CCAGTTGCTCGCGGCTTTTCCGGACACAGTTCCGGATGGTTCAGCCGAGCGCATCAGCA 653

Qy 361 ACCGGAACAATACCGGACAGCGGAACTGCGCTGCGGTTGAGATTAATGACAGCG 420
Db 654 ACCGGAACAATACCGGACAGCGGAACTGCGCTGCGGTTGAGATTAATGACAGCG 713

Qy 421 GTGGCGGCTGGGATATTACGTACGAGGAGCGGTATCCTGGCTGGATGCCGAGAAAT 480
Db 714 GTGGCGGCTGGGATATTACGTACGAGGAGCGGTATCCTGGCTGGATGCCGAGAAAT 773

Qy 481 GGACATGATACCCCGTGAGTTTACCCGGCGGCGGCTCGTTCAATTCACGTTTTTGAAC 540
Db 774 GGACATGATACCCCGTGAGTTTACCCGGCGGCGGCTCGTTCAATTCACGTTTTTGAAC 833

Qy 541 CCGTGAGGAGCGGAGACTGCGGTCGAAATGTGTTTACAGCTGATGAGCGAGTGA 600
Db 834 CCGTGAGGAGCGGAGACTGCGGTCGAAATGTGTTTACAGCTGATGAGCGAGTGA 893

Qy 601 AGATGCTCGACACGCTGACAGAACCGCAGCTAGATTAAACCTAGAAAGATTAATCATTTG 660
Db 894 AGATGCTCGACACGCTGACAGAACCGCAGCTAGATTAAACCTAGAAAGATTAATCATTTG 953

Qy 661 TGACGTACGTTAAAGATAATCATCGTAAAAATTGACGCATGGGATCC 707
Db 954 TGACGTACGTTAAAGATAATCATCGTAAAAATTGACGCATGGGATCC 1000

RESULT 3
US-10-001-189-48/c
; Sequence 48, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
```

```

RESULT 4
US-10-001-189-49/c
; Sequence 49, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; TITLE OF INVENTION: VECTOR PIGGYBAC
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
; LENGTH: 9012
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Description of Artificial Sequence: p(PZ)-Bac-ECFP
; OTHER INFORMATION: sequence
US-10-001-189-49

Query Match 100.0%; Score 707; DB 13; Length 9012;
Best Local Similarity 100.0%; Pred. No. 7.9e-235;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAATCCATCGCTCAATTTTACGACGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
Db 8963 GGAATCCATCGCTCAATTTTACGACGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 8904

Qy 61 ATCATATCGTCCGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGCGGTATCTGGGCA 120
Db 8903 ATCATATCGTCCGGTCTTTTTTCCGGCTCAGTCATCGCCCAAGCTGCGGTATCTGGGCA 8844

Qy 121 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCAGAGTTTGAAGCGGCATGGAAGAGTTTG 180
Db 8943 TCGGGGAGGAAGAAGCCCGTGCCTTTTCCCGCAGAGTTTGAAGCGGCATGGAAGAGTTTG 8784

Qy 181 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAAACGACGTTTACCATGATGATTC 240
Db 8783 CCGAGGATGACTGCTGCTGCATTGACGTTGAGCGAAAAACGACGTTTACCATGATGATTC 8724

Qy 241 GGGNAGTGTGGCCATGACGACGCTTTTAAACGGTGNAACGTGTTTCAGGCCACCTGGGATA 300
Db 8723 GGGNAGTGTGGCCATGACGACGCTTTTAAACGGTGNAACGTGTTTCAGGCCACCTGGGATA 8664

Qy 301 CCAATTTCGTCCGGCTTTTCCGGACACAGTTCGGGATGTTGAGCCCAAGCGCATCAGCA 360
Db 8663 CCAATTTCGTCCGGCTTTTCCGGACACAGTTCGGGATGTTGAGCCCAAGCGCATCAGCA 8604

Qy 361 ACCCGAACAATACCGCGCAACAGCCGGAATCTCCGCTGCGGTGTGCGAGTTAATGACAGCG 420
Db 8603 ACCCGAACAATACCGCGCAACAGCCGGAATCTCCGCTGCGGTGTGCGAGTTAATGACAGCG 8544

Qy 421 GTCCGGCGCTGGGATATTAAGTCAGCGAGACCGGTTATCTTGGCTGATGCCCGAGAAAT 480
Db 8543 GTCCGGCGCTGGGATATTAAGTCAGCGAGACCGGTTATCTTGGCTGATGCCCGAGAAAT 8484

Qy 481 GGCATATGGATAACCCCGTAGTTTACCCGGCGGGCGCGCTCGTTCATTTCAGTTTTTGAAC 540
Db 8483 GGCATATGGATAACCCCGTAGTTTACCCGGCGGGCGCGCTCGTTCATTTCAGTTTTTGAAC 8424

Qy 541 CCGTGGAGGACGGCGAGACTCGCGGTGCAAAATGTGTTTTTACAGCGTGTGAGCAGATGA 600
Db 8423 CCGTGGAGGACGGCGAGACTCGCGGTGCAAAATGTGTTTTTACAGCGTGTGAGCAGATGA 8364

```

Qy 601 AGATGCTCGACACGCTGCAGAACACGCGAGCTAGATTAAACCCCTAGAAAGATAATCATATTG 660
 Db 8363 AGATGCTCGACACGCTGCAGAACACGCGAGCTAGATTAAACCCCTAGAAAGATAATCATATTG 8304
 Qy 661 TGACGTACGTTAAAGATAATCATCGTGAATAAATGACGCGATGGATCC 707
 Db 8303 TGACGTACGTTAAAGATAATCATCGTGAATAAATGACGCGATGGATCC 8257

RESULT 5
 US-10-001-189-50/c
 ; Sequence 50, Application US/10001189
 ; Publication No. US20020173634A1
 ; GENERAL INFORMATION:
 ; APPLICANT: FRASER JR., MALCOLM J.
 ; APPLICANT: LI, XU
 ; APPLICANT: BEAM, TERESA
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
 ; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
 ; FILE REFERENCE: 835910-92098
 ; CURRENT APPLICATION NUMBER: US/10/001,189
 ; CURRENT FILING DATE: 2001-10-30
 ; PRIOR APPLICATION NUMBER: 60/244,984
 ; PRIOR FILING DATE: 2000-11-01
 ; PRIOR APPLICATION NUMBER: 60/244,677
 ; PRIOR FILING DATE: 2000-10-31
 ; NUMBER OF SEQ ID NOS: 70
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 50
 ; LENGTH: 9013
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: p(PZ)-Bac-BGFP
 ; OTHER INFORMATION: sequence
 US-10-001-189-50

Query Match 100.0%; Score 707; DB 13; Length 9013;
 Best Local Similarity 100.0%; Pred. No. 7.9e-235; Mismatches 0; Indels 0; Gaps 0;
 Matches 707; Conservative 0;

Qy 1 GGATCCCATCGCTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
 Db 8964 GGATCCCATCGCTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 8905

Qy 61 ATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120
 Db 8904 ATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 8845

Qy 121 TCGGGGGAAGAAGCCGCTGCTTTTCCCGAGGTTGAAGCGGCATGGAAGAGTTTG 180
 Db 8944 TCGGGGGAAGAAGCCGCTGCTTTTCCCGAGGTTGAAGCGGCATGGAAGAGTTTG 8785

Qy 181 CCGAGGATGACTGCTGCTGCATTTGACGTTGAGCGAAACCGCACGCTTTACCATGATGATTC 240
 Db 8784 CCGAGGATGACTGCTGCTGCATTTGAGCGAAACCGCACGCTTTACCATGATGATTC 8725

Qy 241 GGGAGGTGTGGCCATGACGCGCTTTTAAACGTTGAACTGTTTCAGGCCACCTGGGATA 300
 Db 8724 GGGAGGTGTGGCCATGACGCGCTTTTAAACGTTGAACTGTTTCAGGCCACCTGGGATA 8665

Qy 301 CCAGTTGTCGGGCTTTTCCGGACACAGTTCCGGATGTCAGCCGCGGATGATTC 360
 Db 8664 CCAGTTGTCGGGCTTTTCCGGACACAGTTCCGGATGTCAGCCGCGGATGATTC 8605

Qy 361 ACCGGAACAATACCGGCGACAGCGGAACTGCGTGGCGGTGTGACAGTTAATGACAGCG 420
 Db 8604 ACCGGAACAATACCGGCGACAGCGGAACTGCGTGGCGGTGTGACAGTTAATGACAGCG 8545

Qy 421 GTGGCGGCTGGGATATTAATGTCAGCGAGGACGGGTATCTTGGCTGGATCGCGGAGAAAT 480
 Db 8544 GTGGCGGCTGGGATATTAATGTCAGCGAGGACGGGTATCTTGGCTGGATCGCGGAGAAAT 8485

Qy 481 GGACATGATACCCCGTGAGTTACCCGGCGGCGCTCGTTCAITTCAGCTTTTGAAC 540
 Db 8484 GGACATGATACCCCGTGAGTTACCCGGCGGCGCTCGTTCAITTCAGCTTTTGAAC 8425

Qy 541 CCGTGGAGGACGGGACAGCTCGCGGTGCAAAATGTTTTTACAGGTGATGGAGCAGATGA 600
 Db 8424 CCGTGGAGGACGGGACAGCTCGCGGTGCAAAATGTTTTTACAGGTGATGGAGCAGATGA 8365

Qy 601 AGATGCTCGACACGCTGCAGAACACGCGAGCTAGATTAAACCCCTAGAAAGATAATCATATTG 660
 Db 8364 AGATGCTCGACACGCTGCAGAACACGCGAGCTAGATTAAACCCCTAGAAAGATAATCATATTG 8305

Qy 661 TGACGTACGTTAAAGATAATCATCGTGAATAAATGACGCGATGGATCC 707
 Db 8304 TGACGTACGTTAAAGATAATCATCGTGAATAAATGACGCGATGGATCC 8258

RESULT 6
 US-10-001-189-54
 ; Sequence 54, Application US/10001189
 ; Publication No. US20020173634A1
 ; GENERAL INFORMATION:
 ; APPLICANT: FRASER JR., MALCOLM J.
 ; APPLICANT: LI, XU
 ; APPLICANT: BEAM, TERESA
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
 ; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
 ; FILE REFERENCE: 835910-92098
 ; CURRENT APPLICATION NUMBER: US/10/001,189
 ; CURRENT FILING DATE: 2001-10-30
 ; PRIOR APPLICATION NUMBER: 60/244,984
 ; PRIOR FILING DATE: 2000-11-01
 ; PRIOR APPLICATION NUMBER: 60/244,677
 ; PRIOR FILING DATE: 2000-10-31
 ; NUMBER OF SEQ ID NOS: 70
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 54
 ; LENGTH: 4943
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: PBS-ITR-BCFP
 ; OTHER INFORMATION: sequence
 US-10-001-189-54

Query Match 96.9%; Score 685; DB 13; Length 4943;
 Best Local Similarity 99.2%; Pred. No. 2.7e-227; Mismatches 701; Conservative 0; Indels 6; Gaps 1;

Qy 1 GGATCCCATCGCTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
 Db 719 GGATCCCATCGCTCAATTTTACGCAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 778

Qy 61 ATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120
 Db 779 ATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 838

Qy 121 TCGGGGGAAGAAGCCGCTGCTTTTCCCGAGGTTGAAGCGGCATGGAAGAGTTTG 180
 Db 839 TCGGGGGAAGAAGCCGCTGCTTTTCCCGAGGTTGAAGCGGCATGGAAGAGTTTG 898

Qy 181 CCGAGGATGACTGCTGCTGCATTTGACGTTGAGCGAAACCGCACGCTTTACCATGATGATTC 240
 Db 899 CCGAGGATGACTGCTGCTGCATTTGAGCGAAACCGCACGCTTTACCATGATGATTC 958

Qy 241 GGGAGGTGTGGCCATGACGCGCTTTTAAACGTTGAACTGTTTCAGGCCACCTGGGATA 300
 Db 959 GGGAGGTGTGGCCATGACGCGCTTTTAAACGTTGAACTGTTTCAGGCCACCTGGGATA 1018

Qy 301 CCAGTTGTCGGGCTTTTCCGGACACAGTTCCGGATGTCAGCCGCGGATGATTC 360

Db 1019 CCAGTTCTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAAGCCGAGCGCATCAGCA 1078
Qy 361 ACCGAAACAATACCGCGACAGCGGAACTGCGCGTGCCTGCGGTGTGCAGATTAAATGACAGCG 420
Db 1079 ACCGAAACAATACCGCGACAGCGGAACTGCGCGTGCCTGCGGTGTGCAGATTAAATGACAGCG 1138
Qy 421 GTGCGGCGCTGGGATATTAGTCAAGGAGAGACCGGATATCTGCTGGATCCGCGAGAAAT 480
Db 1139 GTGCGGCGCTGGGATATTAGTCAAGGAGAGACCGGATATCTGCTGGATCCGCGAGAAAT 1198
Qy 481 GGACATCGATACCCCGTGAAGTACCCGCGCGCGCTGCTTCATTTCAGCTTTTGAAC 540
Db 1199 GGACATCGATACCCCGTGAAGTACCCGCGCG-----CTCGTTCAITTCAGCTTTTGAAC 1252
Qy 541 CCGTGGAGGACGGGACAGCTCGCGGTGCAAAATGTGTTTACAGCGTGATGGAGCAGATGA 600
Db 1253 CCGTGGAGGACGGGACAGCTCGCGGTGCAAAATGTGTTTACAGCGTGATGGAGCAGATGA 1312
Qy 601 AGATGCTCGACAGCTGCGAGACACGCGAGCTAGATTAAACCTAGAAAGATAATCATATTG 660
Db 1313 AGATGCTCGACAGCTGCGAGACACGCGAGCTAGATTAAACCTAGAAAGATAATCATATTG 1372
Qy 661 TGACGTACGTTAAAGATAATCATCGCTAAATAATTGACGCGATGGATCC 707
Db 1373 TGACGTACGTTAAAGATAATCATCGCTAAATAATTGACGCGATGGATCC 1419

RESULT 7

US-10-001-189-55
; Sequence 55, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 55
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: PBS-ITR-EGFP
; OTHER INFORMATION: sequence
US-10-001-189-55

Query Match 96.9%; Score 685; DB 13; Length 4944;
Best Local Similarity 99.2%; Pred. No. 2.7e-227;
Matches 701; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

Qy 1 GGATCCCATCGCTCAATTTTACGACAGCTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
Db 719 GGATCCCATCGCTCAATTTTACGACAGCTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 778
Qy 61 ATCATATCTGCGGCTCTTTTCCGGCTAGTCAATGCCCAAGCTGCGCTATCTGGGCA 120
Db 779 ATCATATCTGCGGCTCTTTTCCGGCTAGTCAATGCCCAAGCTGCGCTATCTGGGCA 838
Qy 121 TCGGGAGGAGGAAGCCGCGCTTTTCCCGGAGGTGAAGCGGATGGAAGAGTTTG 180
Db 839 TCGGGAGGAGGAAGCCGCGCTTTTCCCGGAGGTGAAGCGGATGGAAGAGTTTG 898
Qy 181 CCGAGGATGACTGCTGCTGATTTGACGTTGAGCGGAAACGACGCTTTTACCATGATGATTC 240

Db 899 CCAGGATGACTGCTGCTGATTTGAGCGTTGAGCGAAACGACGTTTACCATGATGATTC 958
Qy 241 GGAAGGTGTGGCATGCGCTTTAAACGGTGAACCTGTTTCTAGGCCACCTGGGATA 300
Db 959 GGAAGGTGTGGCATGCGCTTTAAACGGTGAACCTGTTTCTAGGCCACCTGGGATA 1018
Qy 301 CCAAGTTCTGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAAGCCGAGCGCATCAGCA 360
Db 1019 CCAAGTTCTGTCGCGGCTTTTCCGGACACAGTTCCGGATGGTCAAGCCGAGCGCATCAGCA 1078
Qy 361 ACCGAAACAATACCGCGACAGCGGAACTGCGCGTGCCTGCGGTGTGCAGATTAAATGACAGCG 420
Db 1079 ACCGAAACAATACCGCGACAGCGGAACTGCGCGTGCCTGCGGTGTGCAGATTAAATGACAGCG 1138
Qy 421 GTGCGGCGCTGGGATATTAGTCAAGGAGAGACCGGATATCTGCTGGATCCGCGAGAAAT 480
Db 1139 GTGCGGCGCTGGGATATTAGTCAAGGAGAGACCGGATATCTGCTGGATCCGCGAGAAAT 1198
Qy 481 GGACATCGATACCCCGTGAAGTACCCGCGCGCGCTGCTTCATTTCAGCTTTTGAAC 540
Db 1199 GGACATCGATACCCCGTGAAGTACCCGCGCG-----CTCGTTCAITTCAGCTTTTGAAC 1252
Qy 541 CCGTGGAGGACGGGACAGCTCGCGGTGCAAAATGTGTTTACAGCGTGATGGAGCAGATGA 600
Db 1253 CCGTGGAGGACGGGACAGCTCGCGGTGCAAAATGTGTTTACAGCGTGATGGAGCAGATGA 1312
Qy 601 AGATGCTCGACAGCTGCGAGACACGCGAGCTAGATTAAACCTAGAAAGATAATCATATTG 660
Db 1313 AGATGCTCGACAGCTGCGAGACACGCGAGCTAGATTAAACCTAGAAAGATAATCATATTG 1372
Qy 661 TGACGTACGTTAAAGATAATCATCGCTAAATAATTGACGCGATGGATCC 707
Db 1373 TGACGTACGTTAAAGATAATCATCGCTAAATAATTGACGCGATGGATCC 1419

RESULT 8

US-10-001-189-56
; Sequence 56, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPOSITION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: PBS-ITR-EYFP
; OTHER INFORMATION: sequence
US-10-001-189-56

Query Match 96.9%; Score 685; DB 13; Length 4944;
Best Local Similarity 99.2%; Pred. No. 2.7e-227;
Matches 701; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

Qy 1 GGATCCCATCGCTCAATTTTACGACAGCTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
Db 719 GGATCCCATCGCTCAATTTTACGACAGCTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 778


```

Query Match      80.1%; Score 566.2; DB 16; Length 26565;
Best Local Similarity 97.8%; Pred. No. 1.9e-185;
Matches 574; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 45 ATCTAGCTGATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 104
DB 3221 ATCCAGCTGATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 3162

QY 105 TGGCGCTATCTGGGCATCGGGAGGAAGACCGTCCTTTTTCGGGAGGTTGAAGCG 164
DB 3161 TGGCGCTATCTGGGCATCGGGAGGAAGACCGTCCTTTTTCGGGAGGTTGAAGCG 3102

QY 165 GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGCACTTGAACGTTGAGCGAAACGCAAG 224
DB 3101 GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGCACTTGAACGTTGAGCGAAACGCAAG 3042

QY 225 TTACCATGATGATTCGGGAAGTGTGGCCATGACGACCCCTTTTACCGTGAACCTGTCGTT 284
DB 3041 TTACCATGATGATTCGGGAAGTGTGGCCATGACGACCCCTTTTACCGTGAACCTGTCGTT 2982

QY 285 CAGCCACCTTGGGATACAGTTCTGTCGGCTTTTTCGGGACACAGTTCCGGATGTCAGC 344
DB 2981 CAGCCACCTTGGGATACAGTTCTGTCGGCTTTTTCGGGACACAGTTCCGGATGTCAGC 2922

QY 345 CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCCGGAACCTGCGGTGTG 404
DB 2921 CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCCGGAACCTGCGGTGTG 2862

QY 405 CAGATTAATGACAGCGTGGCGCTGGGATATTAACGTCAGGAGGACGGGTATCTCTGC 464
DB 2861 CAGATTAATGACAGCGTGGCGCTGGGATATTAACGTCAGGAGGACGGGTATCTCTGC 2802

QY 465 TGGATCCGCGAGAAATGCAATGATATCCCGTGGATTTACGTCAGGAGGACGGGTATCTCTGC 524
DB 2801 TGGATCCGCGAGAAATGCAATGATATCCCGTGGATTTACGTCAGGAGGACGGGTATCTCTGC 2742

QY 525 ATTACAGCTTTTGAACCCGTTGGAGGACGGGCGAGACTCGCGTGCMAATGTGTTTACAGC 584
DB 2741 ATTACAGCTTTTGAACCCGTTGGAGGACGGGCGAGACTCGCGTGCMAATGTGTTTACAGC 2682

QY 585 GTGATGAGCGATGAAGATGCTCGACACGCTGAGCAACGCGAGCT 631
DB 2681 GTGATGAGCGATGAAGATGCTCGACACGCTGAGCAACGCGAGCT 2635

```

RESULT 11

```

US-10-114-170-72
; Sequence 72, Application US/10114170
; Publication No. US20030023075A1
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
;            Burland, Valerie
;            Perna, Nicole T.
;            Plunkett, Guy
;            Welch, Rod
; TITLE OF INVENTION: No. US20030023075A1el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
;   ADDRESS: Quarles & Brady
;   STREET: 1 South Pinckney Street
;   CITY: Madison
;   STATE: WI
;   COUNTRY: US
;   ZIP: 53701-2113
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/10/114,170
;   FILING DATE: 01-Apr-2002

```

```

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 09/453,702
  FILING DATE: 03-DEC-1999
  APPLICATION NUMBER: 60/110,955
  FILING DATE: 04-DEC-1998
  ATTORNEY/AGENT INFORMATION:
    NAME: Seay, Nicholas J.
    REGISTRATION NUMBER: 27386
    REFERENCE/DOCKET NUMBER: 960296.95017
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (608) 251-5000
    TELEFAX: (608) 251-5166
  INFORMATION FOR SEQ ID NO: 72:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 46819
      TYPE: nucleic acid
      STRANDEDNESS: double
      TOPOLOGY: linear
      MOLECULE TYPE: DNA (genomic)
      SEQUENCE DESCRIPTION: SEQ ID NO: 72:
US-10-114-170-72

Query Match      80.1%; Score 566.2; DB 14; Length 46819;
Best Local Similarity 97.8%; Pred. No. 2.6e-185;
Matches 574; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 45 ATCTAGCTGATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 104
DB 14795 ATCCAGCTGATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 14854

QY 105 TGGCGCTATCTGGGCATCGGGAGGAAGACCGTCCTTTTTCGGGAGGTTGAAGCG 164
DB 14855 TGGCGCTATCTGGGCATCGGGAGGAAGACCGTCCTTTTTCGGGAGGTTGAAGCG 14914

QY 165 GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGCACTTGAACGTTGAGCGAAACGCAAG 224
DB 14915 GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGCACTTGAACGTTGAGCGAAACGCAAG 14974

QY 225 TTACCATGATGATTCGGGAAGTGTGGCCATGACGACCCCTTTTAAACGCTGAACCTGTCGTT 284
DB 14975 TTACCATGATGATTCGGGAAGTGTGGCCATGACGACCCCTTTTAAACGCTGAACCTGTCGTT 15034

QY 285 CAGGCCACCTGGGATACCACTGCTCGCGCTTTTTCGGGACACAGTTCCGGATGTCAGC 344
DB 15035 CAGGCCACCTGGGATACCACTGCTCGCGCTTTTTCGGGACACAGTTCCGGATGTCAGC 15094

QY 345 CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCCGGAACCTGCGGTGTG 404
DB 15095 CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCCGGAACCTGCGGTGTG 15154

QY 405 CAGATTAATGACAGCGTGGCGCTGGGATATTAACGTCAGGAGGACGGGTATCTCTGC 464
DB 15155 CAGATTAATGACAGCGTGGCGCTGGGATATTAACGTCAGGAGGACGGGTATCTCTGC 15214

QY 465 TGGATGCGCAGAAATGGAACATGATGATACCCCGGAGTTACCCCGGCGCGCTCTCTTC 524
DB 15215 TGGATGCGCAGAAATGGAACATGATGATACCCCGGAGTTACCCCGGCGCGCTCTCTTC 15274

QY 525 ATTACAGCTTTTGAACCCGTTGGAGGACGGGCGAGACTCGCGTGCMAATGTGTTTACAGC 584
DB 15275 ATTACAGCTTTTGAACCCGTTGGAGGACGGGCGAGACTCGCGTGCMAATGTGTTTACAGC 15334

QY 585 GTGATGAGCGATGAAGATGCTCGACACGCTGAGCAACGCGAGCT 631
DB 15335 GTGATGAGCGATGAAGATGCTCGACACGCTGAGCAACGCGAGCT 15381

```

RESULT 12

```

US-10-114-170-251/c
; Sequence 251, Application US/10114170
; Publication No. US20030023075A1
; GENERAL INFORMATION:

```

APPLICANT: Blattner, Frederick R.
Burland, Valerie
Perna, Nicole T.
Plunkett, Guy
Welch, Rod
TITLE OF INVENTION: No. US2003023075A1el Sequences of E. coli O157
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/114,170
FILING DATE: 01-Apr-2002
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 09/453,702
FILING DATE: 03-DEC-1999
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 251:
SEQUENCE CHARACTERISTICS:
LENGTH: 22306
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 251:
US-10-114-170-251

Query Match 79.4%; Score 561.4; DB 14; Length 22306;
Best Local Similarity 97.3%; Pred. No. 8.3e-184; Indels 0; Gaps 0;
Matches 571; Conservative 0; Mismatches 16;
45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGCTCAGTCATCGCCCAAGC 104
17193 ATCCAGTTGCATCAGGATCATATCGTCGGGTCTTTTTCGGCTCAGTCATCGCCCAAGC 17134
105 TGGCGCTATCTGGGATCGGGAGAGAAAGCCGTCGCTTTTTCGGGAGAGTTGAAGCG 164
17133 TGGCGCTATCTGGGATCGGGAGAGAAAGCCGTCGCTTTTTCGGGAGAGTTGAAGCG 17074
165 GCATGGAAGAGTTTCCGAGAGATGACCTGCTGCTGATGACCTTGAAGGAAAGCGACG 224
17073 GCATGGAAGAGTTTCCGAGAGATGACCTGCTGCTGATGACCTTGAAGGAAAGCGACG 17014
225 TTATACATGATGATTCGGGAAAGGTGTGGCCATGACGCTTTTAAACGGTGAACCTGTTTCGTT 284
17013 TTATACATGATGATTCGGGAAAGGTGTGGCCATGACGCTTTTAAACGGTGAACCTGTTTCGTT 16954
285 CAGGCCACTGGGATACAGTTTCGTCGGGCTTTTTCGGGACACAGTTCCGGATGTCAGC 344
16953 CAGGCCACTGGGATACAGTTTCGTCGGGCTTTTTCGGGACACAGTTCCGGATGTCAGC 16894
345 CCGAAGCCCATCAGCAACCCGAAACATACCGGACAGCCGAACTGCCCTGCCGCTGTG 404
16893 CCGAAGCCCATCAGCAACCCGAAACATACCGGACAGCCGAACTGCCCTGCCGCTGTG 16834

Qy 405 CAGATTAAATGACAGCGGTGCGGCTGGGATATTACGTACGAGGACGGGATATCTTGGC 464
Db 16833 CAGATTAAATGACAGCGGTGCGGCTGGGATATTACGTACGAGGACGGGATATCTTGGC 16774
Qy 465 TGGATGCCGAGAAATGACATGATACCCGTCGAGTTACCCGCGGCGGCTCGTTC 524
Db 16773 TGGATGCCGAGAAATGACATGATACCCGTCGAGTTACCCGCGGCTCGTTCGTTTC 16714
Qy 525 ATTCAACGTTTGTGAACCCGTCGAGGACGGGACAGACTCGCGTGCAAAATGTGTTTACAGC 584
Db 16713 ATTCAACGTTTGTGAACCCGTCGAGGACGGGACAGACTCGCGTGCAAAATGTGTTTACAGC 16654
Qy 585 GTGATGGAGCAGATGAAGATGCTCGACACGCTCGAGAACGCGAGCT 631
Db 16653 GTGATGGAGCAGATGAAGATGCTCGACACGCTCGAGAACGCGAGCT 16607
RESULT 13
US-10-001-189-41
; Sequence 41, Application US/10001189
; Publication No. US20020173634A1
; GENERAL INFORMATION:
; APPLICANT: FRASER JR., MALCOLM J.
; APPLICANT: LI, XU
; APPLICANT: BEAM, TERESA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSPORTATION USING
; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
; TITLE OF INVENTION: VECTOR PIGGYBAC
; FILE REFERENCE: 835910-92098
; CURRENT APPLICATION NUMBER: US/10/001,189
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/244,984
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,677
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 3662
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: pXL-Bac
; OTHER INFORMATION: sequence
US-10-001-189-41

Query Match 73.4%; Score 519.2; DB 13; Length 3662;
Best Local Similarity 98.5%; Pred. No. 1.7e-169; Indels 0; Gaps 0;
Matches 524; Conservative 0; Mismatches 8;
2 GATCCCATGGCTCAATTTTACGAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGGA 61
983 GATCCCATGGCTCAATTTTACGAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGGA 1042
62 TCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGCTGCGGCTATCTGGGAT 121
1043 TCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGCTGCGGCTATCTGGGAT 1102
122 CGGGAGAGAGAGCCGTCGCTTTTCCCGGAGTTGAGCGGCATGGAAGAGTTTGC 181
1103 CGGGAGAGAGAGCCGTCGCTTTTCCCGGAGTTGAGCGGCATGGAAGAGTTTGC 1162
182 CGAGGATGACTGCTGCTGCATTGACGTTGAGCGGAAACCGACGTTTACCATGATGATTCG 241
1163 CGAGGATGACTGCTGCTGCATTGACGTTGAGCGGAAACCGACGTTTACCATGATGATTCG 1222
242 GGAAGGTGTGGCCATGACGCTTTTAAACGGTGAACCTGTTCCGTCAGGACCTGGGATAC 301
1223 GGAAGGTGTGGCCATGACGCTTTTAAACGGTGAACCTGTTCCGTCAGGACCTGGGATAC 1282
302 CAGTTCTCGCGGCTTTTCCGGACACAGTTCCGGATGTCAGCCCGGAGCGCATCAGCAA 361
1283 CAGTTCTCGCGGCTTTTCCGGACACAGTTCCGGATGTCAGCCCGGAGCGCATCAGCAA 1342

Qy 362 CCCGAAATACCCGGGACAGCCGGAATCTCCGCTGCCGGTGTGCAGATTAATGACACCGG 421
 Db 1343 CCCGAAATACCCGGGACAGCCGGAATCTCCGCTGCCGGTGTGCAGATTAATGACACCGG 1402
 Qy 422 TCGCGGCTGGATATTACGTCAGCGAGGACGGGTATCTTGGCTGGATGCGGAGAAATG 481
 Db 1403 TCGCGGCTGGATATTACGTCAGCGAGGACGGGTATCTTGGCTGGATGCGGAGAAATG 1462
 Qy 482 GACATGATACCCCGTGGATTTACCCGGCGGCGGCTCTGTTCAATTCACGTT 533
 Db 1463 GACATGATACCCCGTGGATTTACCCGGCGGCGGCTCTGCGTGAATCATGTT 1514

RESULT 14
 US-10-001-189-53/c
 ; Sequence 53, Application US/10001189
 ; Publication No. US20020173634A1
 ; GENERAL INFORMATION:
 ; APPLICANT: FRASER JR., MALCOLM J.
 ; APPLICANT: LI, XU
 ; APPLICANT: BEAM, TERESA
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSDUCTION USING
 ; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
 ; TITLE OF INVENTION: VECTOR PIGGYBAC
 ; FILE REFERENCE: 835910-92098
 ; CURRENT APPLICATION NUMBER: US/10/001.189
 ; PRIOR FILING DATE: 2001-10-30
 ; PRIOR APPLICATION NUMBER: 60/244,984
 ; PRIOR FILING DATE: 2000-11-01
 ; PRIOR APPLICATION NUMBER: 60/244,677
 ; PRIOR FILING DATE: 2000-10-31
 ; NUMBER OF SEQ ID NOS: 70
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 53
 ; LENGTH: 4941
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: pXL-Bac-ECFP
 ; OTHER INFORMATION: sequence
 US-10-001-189-53

Query Match 72.3%; Score 511; DB 13; Length 4941;
 Best Local Similarity 100.0%; Pred. No. 1.5e-166;
 Matches 511; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 7 CATCGCTCAATTTTACGACACTATCTTTCTAGGGTTAATCTAGCTGCATCAGATCATA 66
 Db 1129 CATCGCTCAATTTTACGACACTATCTTTCTAGGGTTAATCTAGCTGCATCAGATCATA 1070
 Qy 67 TCGTCCGGTCTTTTTCGGCTCAGTCATCGCCCAAGCTGGCGTATCTGGGATCGGG 126
 Db 1069 TCGTCCGGTCTTTTTCGGCTCAGTCATCGCCCAAGCTGGCGTATCTGGGATCGGG 1010
 Qy 127 AGGAAGAGCCCGTGCCTTTTCCGCGAGGTGAAGCGGATGGAAGAGTTTCCGAGG 186
 Db 1009 AGGAAGAGCCCGTGCCTTTTCCGCGAGGTGAAGCGGATGGAAGAGTTTCCGAGG 950
 Qy 187 ATGACTCTCTGCTGATTTGAGCGGAAACGACAGTTTACCATGATGATTCGGGAAG 246
 Db 949 ATGACTCTCTGCTGATTTGAGCGGAAACGACAGTTTACCATGATGATTCGGGAAG 890
 Qy 247 GTGTGGCCATGACGCGCTTTTAAACGTTGAACTGTTTTCAGGCCACCTGGGATACCAATT 306
 Db 889 GTGTGGCCATGACGCGCTTTTAAACGTTGAACTGTTTTCAGGCCACCTGGGATACCAATT 830
 Qy 307 CGTCCGGCTTTTCCGACACAGTTCCGGATGTTCCGGTGTGACGCGGATCAGCAACCCGA 366
 Db 829 CGTCCGGCTTTTCCGACACAGTTCCGGATGTTCCGGTGTGACGCGGATCAGCAACCCGA 770
 Qy 367 ACATACCCGGGACAGCCGGAATCTCCGCTGCCGGTGTGAGATTAATGACAGCGGTGCGG 426
 Db 367 ACATACCCGGGACAGCCGGAATCTCCGCTGCCGGTGTGAGATTAATGACAGCGGTGCGG

Db 769 ACAATACCCGGGACAGCCGGAATCTGCCGTGCCGGTGTGCAGATTAATGACAGCGGTGCGG 710
 Qy 427 CGCTGGGATATTACGTCAGCGAGGACGGGTATCTTGGCTGGATGCGGACAAATGGACAT 486
 Db 709 CGCTGGGATATTACGTCAGCGAGGACGGGTATCTTGGCTGGATGCGGACAAATGGACAT 650
 Qy 487 GATATACCCCGTGGATTTACCCGGCGGCGGCGC 517
 Db 649 GATATACCCCGTGGATTTACCCGGCGGCGGCGC 619

RESULT 15
 US-10-001-189-51/c
 ; Sequence 51, Application US/10001189
 ; Publication No. US20020173634A1
 ; GENERAL INFORMATION:
 ; APPLICANT: FRASER JR., MALCOLM J.
 ; APPLICANT: LI, XU
 ; APPLICANT: BEAM, TERESA
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TRANSDUCTION USING
 ; TITLE OF INVENTION: MINIMAL SEGMENTS OF THE EUKARYOTIC TRANSFORMATION
 ; TITLE OF INVENTION: VECTOR PIGGYBAC
 ; FILE REFERENCE: 835910-92098
 ; CURRENT APPLICATION NUMBER: US/10/001.189
 ; PRIOR FILING DATE: 2001-10-30
 ; PRIOR APPLICATION NUMBER: 60/244,984
 ; PRIOR FILING DATE: 2000-11-01
 ; PRIOR APPLICATION NUMBER: 60/244,677
 ; PRIOR FILING DATE: 2000-10-31
 ; NUMBER OF SEQ ID NOS: 70
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 51
 ; LENGTH: 4951
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: pXL-Bac-BYFP
 ; OTHER INFORMATION: sequence
 US-10-001-189-51
 Query Match 72.3%; Score 511; DB 13; Length 4951;
 Best Local Similarity 100.0%; Pred. No. 1.5e-166;
 Matches 511; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 7 CATCGCTCAATTTTACGACACTATCTTTCTAGGGTTAATCTAGCTGCATCAGATCATA 66
 Db 1129 CATCGCTCAATTTTACGACACTATCTTTCTAGGGTTAATCTAGCTGCATCAGATCATA 1070
 Qy 67 TCGTCCGGTCTTTTTCGGCTCAGTCATCGCCCAAGCTGGCGTATCTGGGATCGGG 126
 Db 1069 TCGTCCGGTCTTTTTCGGCTCAGTCATCGCCCAAGCTGGCGTATCTGGGATCGGG 1010
 Qy 127 AGGAAGAGCCCGTGCCTTTTCCGCGAGGTGAAGCGGATGGAAGAGTTTCCGAGG 186
 Db 1009 AGGAAGAGCCCGTGCCTTTTCCGCGAGGTGAAGCGGATGGAAGAGTTTCCGAGG 950
 Qy 187 ATGACTCTCTGCTGATTTGAGCGGAAACGACAGTTTACCATGATGATTCGGGAAG 246
 Db 949 ATGACTCTCTGCTGATTTGAGCGGAAACGACAGTTTACCATGATGATTCGGGAAG 890
 Qy 247 GTGTGGCCATGACGCGCTTTTAAACGTTGAACTGTTTTCAGGCCACCTGGGATACCAATT 306
 Db 889 GTGTGGCCATGACGCGCTTTTAAACGTTGAACTGTTTTCAGGCCACCTGGGATACCAATT 830
 Qy 307 CGTCCGGCTTTTCCGACACAGTTCCGGATGTTCCGGTGTGACGCGGATCAGCAACCCGA 366
 Db 829 CGTCCGGCTTTTCCGACACAGTTCCGGATGTTCCGGTGTGACGCGGATCAGCAACCCGA 770
 Qy 367 ACATACCCGGGACAGCCGGAATCTCCGCTGCCGGTGTGAGATTAATGACAGCGGTGCGG 426
 Db 769 ACAATACCCGGGACAGCCGGAATCTCCGCTGCCGGTGTGAGATTAATGACAGCGGTGCGG 710
 Qy 427 CGCTGGGATATTACGTCAGCGAGGACGGGTATCTTGGCTGGATGCGGACAAATGGACAT 486

Db	709	CGCTGGGATATTACGTACAGCAGGACGGGTATCCTGGCTGGATGCCGCAAAATGGACAT	650
Qy	487	GGATACCCCGTGAGTTACCCGGCGGCGCGC	517
Db	649	GGATACCCCGTGAGTTACCCGGCGGCGCGC	619

Search completed: October 13, 2005, 13:50:38
 Job time : 5850 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 12, 2005, 13:49:25 ; Search time 26459 Seconds
(without alignments)
1294.752 Million cell updates/sec

Title: US-10-826-523-40
Perfect score: 707
Sequence: 1 ggatccatgcgtcaatttt.....aaattgacgcgtggatcc 707

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	585.4	82.8	20067	12 XXU02425	U02425 Cloning vec
2	585.4	82.8	20125	12 XXU02426	U02426 Cloning vec
3	585.4	82.8	42529	12 CVU39284	U39284 Cloning vec
4	585.4	82.8	42530	12 CVU39285	U39285 Cloning vec
5	585.4	82.8	42531	12 CVU39286	U39286 Cloning vec
6	585.4	82.8	42704	12 CVU37692	U37692 Cloning vec
7	585.4	82.8	48502	7 LAMCG	J02459 Bacterioph
8	574.6	81.3	5969	12 AF434923	AF434923 Expressio
9	574.6	81.3	6182	12 AF434925	AF434925 Expressio
10	574.6	81.3	6182	12 AF434926	AF434926 Expressio
11	574.6	81.3	6182	12 AF434927	AF434927 Expressio
12	574.6	81.3	6182	12 AF434931	AF434931 Expressio
13	574.6	81.3	6182	12 AF434932	AF434932 Expressio
14	574.6	81.3	7421	12 AF434924	AF434924 Expressio
15	574.6	81.3	7634	12 AF434928	AF434928 Expressio
16	574.6	81.3	7634	12 AF434929	AF434929 Expressio
17	574.6	81.3	7634	12 AF434930	AF434930 Expressio
18	574.6	81.3	7634	12 AF434933	AF434933 Expressio
19	574.6	81.3	7634	12 AF434934	AF434934 Expressio

20	566.2	80.1	14164	1 AE005330	AE005330 Escherich
21	566.2	80.1	26565	6 AX702505	AX702505 Sequence
22	566.2	80.1	46819	6 AR204176	AR204176 Sequence
23	566.2	80.1	305325	1 AE016765	AE016765 Escherich
24	564.6	79.9	46897	6 BD184772	BD184772 Nucleic a
25	564.6	79.9	222605	1 AP002555	AP002555 Escherich
26	561.4	79.4	22306	6 AR204355	AR204355 Sequence
27	561.4	79.4	91772	6 BD184765	BD184765 Nucleic a
28	561.4	79.4	286485	1 AP002557	AP002557 Escherich
29	407	57.6	46375	7 AF064539	AF064539 Bacteriop
30	395.8	56.0	64137	2 AC101160	AC101160 Mus muscu
31	395.8	56.0	179840	2 AC068458	AC068458 Homo sapi
32	395.8	56.0	179871	2 AC024602	AC024602 Homo sapi
33	395.8	56.0	194028	2 AC111077	AC111077 Homo sapi
34	395.8	56.0	206486	2 AC074218	AC074218 Mus muscu
35	362.2	51.2	228516	2 AC008594	AC008594 Homo sapi
36	343.2	48.5	68028	2 AC101473	AC101473 Mus muscu
37	329.8	46.6	66450	2 AC060821	AC060821 Homo sapi
38	320.4	45.3	63579	2 AC100018	AC100018 Mus muscu
39	320.2	45.3	63690	2 AC125419	AC125419 Homo sapi
40	231.8	32.8	29466	1 AE008818	AE008818 Salmonell
41	223	31.5	8134	7 PH2HEADTL	M81255 Bacterioph
42	223	31.5	13264	1 AE005368	AE005368 Escherich
43	223	31.5	13386	1 AE015100	AE015100 Shigella
44	223	31.5	34063	6 AR204200	AR204200 Sequence
45	223	31.5	44029	6 BD184774	BD184774 Nucleic a

ALIGNMENTS

RESULT 1
XXU02425
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
ORIGIN
Query Match
Best Local Similarity
Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XXU02425 20067 bp DNA linear SYN 29-JAN-1997
Cloning vector lambda EMBL3, left arm.
U02425
U02425.1 GI:413791
Cloning vector lambda EMBL3
Cloning vector lambda EMBL3
other sequences; artificial sequences; vectors.
1 (bases 1 to 20067)
Frischauf,A.M., Lehrsch,H., Poustka,A. and Murray,N.
Lambda replacement vectors carrying polylinker sequences
J. Mol. Biol. 170 (4), 827-842 (1983)
84064856
6315951
2 (bases 1 to 20067)
Kitts,P.A.
CLONTECH Vectors On Disc version 1.3
Unpublished
3 (bases 1 to 20067)
Kitts,P.A.
Direct Submission
Submitted (07-OCT-1993) Paul A. Kitts, CLONTECH Laboratories, Inc.,
1020 East Meadow Circle, Palo Alto, CA 94303, USA
This sequence has been compiled from information in the sequence
databases, published literature and other sources. If you suspect
there is an error in this sequence, please contact CLONTECH's
Technical Service Department at (415) 424-8222 or (800) 662-2566,
extension 3 or E-mail TECH@CLONTECH.COM.
1. .20067
Location/Qualifiers
/organism="Cloning vector lambda EMBL3"
/mol_type="genomic DNA"
/db_xref="taxon:11783"

QY

45 ATCTAGTCGTCATCAGGATCATATCGTCGGGTCTTTTTCGCGGTCTAGTCATCGCCCAAGC 104

```
|||||
Db 3055 ATCCAGCTGCATCAGGATCATATCGTCGGGTCCTTTTCCGGCTCAGTCATCGCCCAAGC 3114
Qy 105 TGGCGCTATCTGGGATCGGGAGAGAAAGACCCGTCGCTTTTCCCGAGAGTTGAAGCG 164
Db 3115 TGGCGCTATCTGGGATCGGGAGAGAAAGACCCGTCGCTTTTCCCGAGAGTTGAAGCG 3174
Qy 165 GCATGGAAAGAGTTTGGCGAGGATGACTGCTGCTGCAATGACGTTGAGCGAAACGACG 224
Db 3175 GCATGGAAAGAGTTTGGCGAGGATGACTGCTGCTGCAATGACGTTGAGCGAAACGACG 3234
Qy 225 TTTACCATGATGATTTCCGGAAGAGTGTGCCATGCAAGCTTTTAAACGGTGAACCTGTTGTT 284
Db 3235 TTTACCATGATGATTTCCGGAAGAGTGTGCCATGCAAGCTTTTAAACGGTGAACCTGTTGTT 3294
Qy 285 CAGGCCACCTGGGATACCAAGTTCTGTCGGGCTTTTCCGGAACAGATTCGGATCGGTGACG 344
Db 3295 CAGGCCACCTGGGATACCAAGTTCTGTCGGGCTTTTCCGGAACAGATTCGGATCGGTGACG 3354
Qy 345 CCGAAGCGCATCAGCAACCCGAAACAAATACCGGCGACAGCCGGAACCTGCGGTGCGGTGTG 404
Db 3355 CCGAAGCGCATCAGCAACCCGAAACAAATACCGGCGACAGCCGGAACCTGCGGTGCGGTGTG 3414
Qy 405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTACGAGAGAGCGGTATCCTGGC 464
Db 3415 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTACGAGAGAGCGGTATCCTGGC 3474
Qy 465 TGGATGCGCGAGAAATGACATGATACCCGCTGAGTTACCCGCGGCGCGCTCGTTTC 524
Db 3475 TGGATGCGCGAGAAATGACATGATACCCGCTGAGTTACCCGCGGCGCGCTCGTTTC 3534
Qy 525 ATTACAGCTTTTGAACCCGTCGAGAGAGCGGAGACTCGCGTGCAAAATGCTTTTACAGC 584
Db 3535 ATTACAGCTTTTGAACCCGTCGAGAGAGCGGAGACTCGCGTGCAAAATGCTTTTACAGC 3594
Qy 585 GTGATGAGCAGATGAAGATGCTGCACACGCTGCAGAACACGCGACT 631
Db 3595 GTGATGAGCAGATGAAGATGCTGCACACGCTGCAGAACACGCGACT 3641
```

```
RESULT 2
XXU02426
LOCUS Cloning vector lambda EMBL3 SP6/T7, left arm. linear SYN 29-JAN-1997
ACCESSION U02426
VERSION U02426.1 GI:413792
KEYWORDS Cloning vector lambda EMBL3 SP6/T7
SOURCE Cloning vector lambda EMBL3 SP6/T7
ORGANISM other sequences; artificial sequences; vectors.
REFERENCE 1 (bases 1 to 20125)
AUTHORS Kitts,P.A.
TITLE CLONTECH Vectors On Disc version 1.3
JOURNAL Unpublished
AUTHORS Kitts,P.A.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1993) Paul A. Kitts, CLONTECH Laboratories, Inc., 1020 East Meadow Circle, Palo Alto, CA 94303, USA
COMMENT This sequence has been compiled from information in the sequence databases, published literature and other sources. If you suspect there is an error in this sequence, please contact CLONTECH's Technical Service Department at (415) 424-8222 or (800) 662-2566, extension 3 or E-mail TECH@CLONTECH.COM.
```

```
FEATURES
source
1..20125
/organism="Cloning vector lambda EMBL3 SP6/T7"
/mol_type="genomic DNA"
/db_xref="taxon:31784"
```

ORIGIN

```
Query Match 82.8%; Score 585.4; DB 12; Length 20125;
Best Local Similarity 99.8%; Pred. No. 2.3e-148;
```

```
Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCCTTTTCCGGCTCAGTCATCGCCCAAGC 104
Db 3055 ATCCAGCTGCATCAGGATCATATCGTCGGGTCCTTTTCCGGCTCAGTCATCGCCCAAGC 3114
Qy 105 TGGCGCTATCTGGGATCGGGAGAGAAAGACCCGTCGCTTTTCCCGAGAGTTGAAGCG 164
Db 3115 TGGCGCTATCTGGGATCGGGAGAGAAAGACCCGTCGCTTTTCCCGAGAGTTGAAGCG 3174
Qy 165 GCATGGAAAGAGTTTGGCGAGGATGACTGCTGCTGCAATGACGTTGAGCGAAACGACG 224
Db 3175 GCATGGAAAGAGTTTGGCGAGGATGACTGCTGCTGCAATGACGTTGAGCGAAACGACG 3234
Qy 225 TTTACCATGATGATTTCCGGAAGAGTGTGCCATGCAAGCTTTTAAACGGTGAACCTGTT 284
Db 3235 TTTACCATGATGATTTCCGGAAGAGTGTGCCATGCAAGCTTTTAAACGGTGAACCTGTT 3294
Qy 285 CAGGCCACCTGGGATACCAAGTTCTGTCGGGCTTTTCCGGAACAGATTCGGATCGGTGACG 344
Db 3295 CAGGCCACCTGGGATACCAAGTTCTGTCGGGCTTTTCCGGAACAGATTCGGATCGGTGACG 3354
Qy 345 CCGAAGCGCATCAGCAACCCGAAACAAATACCGGCGACAGCCGGAACCTGCGGTGCGGTGTG 404
Db 3355 CCGAAGCGCATCAGCAACCCGAAACAAATACCGGCGACAGCCGGAACCTGCGGTGCGGTGTG 3414
Qy 405 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTACGAGAGAGCGGTATCCTGGC 464
Db 3415 CAGATTAATGACAGCGGTGCGGCGCTGGGATATTACGTACGAGAGAGCGGTATCCTGGC 3474
Qy 465 TGGATGCGCGAGAAATGACATGATACCCGCTGAGTTACCCGCGGCGCGCTCGTTTC 524
Db 3475 TGGATGCGCGAGAAATGACATGATACCCGCTGAGTTACCCGCGGCGCGCTCGTTTC 3534
Qy 525 ATTACAGCTTTTGAACCCGTCGAGAGAGCGGAGACTCGCGTGCAAAATGCTTTTACAGC 584
Db 3535 ATTACAGCTTTTGAACCCGTCGAGAGAGCGGAGACTCGCGTGCAAAATGCTTTTACAGC 3594
Qy 585 GTGATGAGCAGATGAAGATGCTGCACACGCTGCAGAACACGCGACT 631
Db 3595 GTGATGAGCAGATGAAGATGCTGCACACGCTGCAGAACACGCGACT 3641
```

```
RESULT 3
CVU39284
LOCUS Cloning vector TLF97-1, lambda phage lacZ translational fusion
DEFINITION vector, complete sequence.
ACCESSION U39284
VERSION U39284.1 GI:1066304
KEYWORDS Cloning vector TLF97-1
SOURCE Cloning vector TLF97-1
ORGANISM other sequences; artificial sequences; vectors.
REFERENCE 1 (bases 1 to 42529)
AUTHORS St Pierre,R. and Linn,T.
TITLE A refined vector system for the in vitro construction of single-copy transcriptional or translational fusions to lacZ
```

```
JOURNAL Gene 169 (1), 65-68 (1996)
MEDLINE 96186904
PUBMED 8635751
REFERENCE 2 (bases 1 to 42529)
AUTHORS StPierre,R.
TITLE Direct Submission
JOURNAL Submitted (24-OCT-1995) Thomas Linn, Microbiology and Immunology, University of Western Ontario, London, Ontario N6A5C1, Canada
FEATURES
source
1..42529
/organism="Cloning vector TLF97-1"
/mol_type="genomic DNA"
/db_xref="taxon:43840"
complement(19404..19489)
/note="rrnBT2 terminator"
terminator
```

[illegible]

gene	complement (19507. .20760) /gene="lacY"	Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
CDS	complement (19507. .20760) /gene="lacY" /codon_start=1 /transl_table=11 /product="lactose permease" /protein_id="AAC53651.1" /db_xref="GI:1066309"	45 ATCTAGCTGCATCAGGATCATATCGTCGGGCTCTTTTTCGGGCTCAGTATCGCCCAAGC 104 3055 ATCAGCTGCATCAGGATCATATCGTCGGGCTCTTTTTCGGGCTCAGTATCGCCCAAGC 3114 105 TGGCGCTATCTGGGCATCGGGGAGGAAGACCCTGCTCTTTTTCGGGAGGTTGAAGCG 164 3115 TGGCGCTATCTGGGCATCGGGGAGGAAGACCCTGCTCTTTTTCGGGAGGTTGAAGCG 3174 165 GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGCAATGAGCTTGAGCGAAAGCGACG 224 3175 GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGCAATGAGCTTGAGCGAAAGCGACG 3234 225 TTTACCATGATGATTCGGGAAGGTGTGCCCATGCACGCTTTTAAACGGTGAACTGTTCTGT 284 3235 TTTACCATGATGATTCGGGAAGGTGTGCCCATGCACGCTTTTAAACGGTGAACTGTTCTGT 3294 285 CAGGCCACCTGGGATACCAAGTTCTGCGCGCTTTTCCGGACACAGATTCGGAATGCTCAGC 344 3295 CAGGCCACCTGGGATACCAAGTTCTGCGCGCTTTTCCGGACACAGATTCGGAATGCTCAGC 3354 345 CCGAAGCGCATCAGCAACCCGCAACAATACCGCGCACAGCCGGAATCCCGTCCCGGTGTG 404 3355 CCGAAGCGCATCAGCAACCCGCAACAATACCGCGCACAGCCGGAATCCCGTCCCGGTGTG 3414 405 CAGATTAATGACAGCGGTGCGCGCTGCGGATATTACCTCAGCAGGAGCGGATATCTTGGC 464 3415 CAGATTAATGACAGCGGTGCGCGCTGCGGATATTACCTCAGCAGGAGCGGATATCTTGGC 3474 465 TGGATGCGCGAGAAATGGACATGATACCCCGTGAATACCCCGGCGGCGCGCTCGTTC 524 3475 TGGATGCGCGAGAAATGGACATGATACCCCGTGAATACCCCGGCGGCGCGCTCGTTC 3534 525 ATTACGCTTTTGAACCGGTGAGGACCGGCGAGCTCGCGTGCAGATGCTTTTACAGC 584 3535 ATTACGCTTTTGAACCGGTGAGGACCGGCGAGCTCGCGTGCAGATGCTTTTACAGC 3594 585 GTGATGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCGAGCT 631 3595 GTGATGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCGAGCT 3641
misc_feature	CUU39286 42531 bp DNA linear SYN 13-APR-1996 Cloning vector TLF97-3, phage lambda lacZ translational fusion vector, complete sequence.	
terminator	ACCESSION U39286 VERSION U39286.1 GI:1066312 KEYWORDS SOURCE ORGANISM REFERENCE AUTHORS TITLE JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE JOURNAL FEATURES source	
gene	24476. .25666 /gene="tetr"	
CDS	24476. .25666 /gene="tetr" /codon_start=1 /transl_table=11 /product="tetracycline resistance protein" /protein_id="AAC53653.1" /db_xref="GI:1066311"	
misc_feature	33603. .35451 /note="imm21 substitution"	
ORIGIN		
Query Match	82.8%; Score 585.4; DB 12; Length 42530;	
Best Local Similarity	99.8%; Pred. No. 2.5e-148;	

Query Match 82.8%; score 585.4; DB 12; Length 42531;
Best Local Similarity 99.8%; Pred. No. 2.5e-148;

959843
13 (bases 35578 to 35667)
Humayun, Z., Jeffrey, A. and Ptashne, M.
Completed DNA sequences and organization of repressor-binding sites
in the operators of phage lambda
J. Mol. Biol. 112 (2), 265-277 (1977)
77209970
875019
14 (bases 38610 to 38732)
Scherer, G., Hobom, G. and Kossel, H.
DNA base sequence of the po promoter region of phage lambda
Nature 265 (5590), 117-121 (1977)
77100320
834253
15 (bases 38041 to 38241)
Roberts, T.M., Shimatake, H., Brady, C. and Rosenberg, M.
Sequence of Cro gene of bacteriophage lambda
Nature 270 (5634), 274-275 (1977)
78071724
593347
16 (bases 27616 to 28935)
Davies, R.W., Schreier, P.H. and Buchel, D.E.
Nucleotide sequence of the attachment site of coliphage lambda
Nature 270 (5639), 757-760 (1977)
78071823
593399
17 (bases 37206 to 37263; 37914 to 37970)
Humayun, Z.
DNA sequence at the end of the CI gene in bacteriophage lambda
Nucleic Acids Res. 4 (7), 2137-2143 (1977)
78011659
909767
18 (bases 27617 to 27934)
Landy, A. and Ross, W.
Viral integration and excision: structure of the lambda att sites
Science 197 (4309), 1147-1160 (1977)
77258934
331474
19 (bases 39062 to 39170)
Denniston-Thompson, K., Moore, D.D., Kruger, K.E., Furth, M.E. and
Blattner, F.R.
Physical structure of the replication origin of bacteriophage
lambda
Science 198 (4321), 1051-1056 (1977)
78054731
929187
20 (bases 4467 to 44807)
Sklar, J.L.
Structure and function of two regions of DNA controlling the
synthesis of prokaryotic RNAs
Thesis (1977)
21 (sites)
Adhya, S. and Gottesman, M.
Control of transcription termination
Annu. Rev. Biochem. 47, 967-996 (1978)
78234064
354508
22 (bases 13 to 72; 48391 to 48502)
Nichols, B.P. and Donelson, J.E.
178-Nucleotide sequence surrounding the cos site of bacteriophage
lambda DNA
J. Virol. 26 (2), 429-434 (1978)
78197067
666898
23 (bases 37938 to 38016; 35589 to 35666)
Flashman, S.M.
Mutational analysis of the operators of bacteriophage lambda
Mol. Gen. Genet. 166 (1), 61-73 (1978)
79114073
368570
24 (bases 37990 to 38982)
Schwarz, E., Scherer, G., Hobom, G. and Kossel, H.
Nucleotide sequence of cro, cII and part of the O gene in phage

DRPELNEAI PNDERDTPMPVAMATTLRLKLTGLASRQOLIDWMEADKVAGPL
LRSLPAGWFIADKSGAGRSRGIITAAALPDGKPRIVVIYTTGSOATMDERNQIA
EIGASLIKHW
5316..5735
/note="ColE1"

rep_origin

ORIGIN

Query Match 81.3%; Score 574.6; DB 12; Length 5969;
Best Local Similarity 99.3%; Pred. No. 1.8e-145;
Matches 577; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 45 ATCTAGCTGATCAGGATCATATCGTCGGGCTCTTTTCCGGCTCAGTCATCGCCCAAGC 104
|||
Db 3508 ATCCAGCTGATCAGGATCATATCGTCGGGCTCTTTTCCGGCTCAGTCATCGCCCAAGC 3567
|||
Qy 105 TGGCGCTATCTGGGCATCGGGAGGAAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
|||
Db 3568 TGGCGCTATCTGGGCATCGGGAGGAAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 3627
|||
Qy 165 GCATGGAAGAGTTTCCGAGGATGACTGCTGTGCATGACGTTGAGCGAAACGCAACG 224
|||
Db 3628 GCATGGAAGAGTTTCCGAGGATGACTGCTGTGCATGACGTTGAGCGAAACGCAACG 3687
|||
Qy 225 TTTACCATGATGTTCCGAGGATGACTGCTGTGCATGACGTTTAAACGTTGACTTTCGTT 284
|||
Db 3688 TTTACCATGATGATTCGGGAAGGTTGCGCATGACGCTTTTAAACGTTGACTTTCGTT 3747
|||
Qy 285 CAGGCCACCTGGGATACAGTTCGTCGGGCTTTCCGGACACAGTTCGGATGGTGTCAGC 344
|||
Db 3748 CAGGCCACCTGGGATACAGTTCGTCGGGCTTTCCGGACACAGTTCGGATGGTGTCAGC 3807
|||
Qy 345 CCGAAGCGCATGACAAACCCGAAATACCGGGACAGCCGGAATCGCCGTGCGGTTG 404
|||
Db 3808 CCGAAGCGCATGACAAACCCGAAATACCGGGACAGCCGGAATCGCCGTGCGGTTG 3867
|||
Qy 405 CAGATTAATGACAGCGTGGCGCTGGGATATACGTCAGCGAGGAGCGGTATCTTCGC 464
|||
Db 3868 CAGATTAATGACAGCGTGGCGCTGGGATATACGTCAGCGAGGAGCGGTATCTTCGC 3927
|||
Qy 465 TGGATGCGCAGAAATGACATGATACCCCGTGAGTTACCCGCGCGCGCTCGTTC 524
|||
Db 3928 TGGATGCGCAGAAATGACATGATACCCCGTGAGTTACCCGCGCGCGCTCGTTC 3987
|||
Qy 525 ATTACGCTTTTGAACCCGTGGAGGACGCGCAGACTCGCGTGCAATGTGTTTACAGC 584
|||
Db 3988 ATTACGCTTTTGAACCCGTGGAGGACGCGCAGACTCGCGTGCAATGTGTTTACAGC 4047
|||
Qy 585 GTGATGAGCAGATGAGATGCTCGACACGCTGCAGAACAC 625
|||
Db 4048 GTGATGAGCAGATGAGATGCTCGACACGCTGCATGAAC 4088
|||

RESULT 9
AF434925
LOCUS
DEFINITION
Expression vector 409-FOR, complete sequence.
AF434925
ACCESSION
AF434925.1 GI:16904158
KEYWORDS
SOURCE
ORGANISM
Expression vector 409-FOR
Expression vector 409-FOR
other sequences; artificial sequences; vectors.
REFERENCE
1 (bases 1 to 6182)
Zieler,H. and Huynh,C.Q.
Intron-dependent stimulation of marker gene expression in cultured
insect cells
JOURNAL
Insect Mol. Biol. 11 (1), 87-95 (2002)
MEDLINE
21830665
PUBMED
11841506
REFERENCE
2 (bases 1 to 6182)
Zieler,H. and Huynh,C.Q.
Direct Submission
AUTHORS
Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of

Parasitic Diseases, National Institutes of Health, 4 Center Drive
MSC 0425, Bethesda, MD 20892-0425, USA
Location/Qualifiers
1..6182
/organism="Expression vector 409-FOR"
/mol_type="genomic DNA"
/db_xref="taxon:1176967"
/notes="insect-specific"
97..614
/notes="Autographa californica nuclear polyhedrosis virus
hr5 enhancer"
615..1201
/note="Autographa californica nuclear polyhedrosis virus
immediate early promoter"
1266..1398
/notes="artificial intron"
1541..3193
/codon_start=1
/product="firefly luciferase"
/protein_id="AAL30782.1"
/db_xref="GI:16904159"
/translation="MEDAKNIKKGPAPFYPLEDTAGTEQLHKAMRYALVPGTIAFTD
AHLEVNITVAEYPEMSVRLAEAMKRYGLNTNHRIVVCSNSLOFEMVLGALFIGVAV
APANDLYNERELLNSMNI SQPTVVVSKGKLOKILNVOKKLPLOKIIIMDSKTDYOG
FQSMYTFVSHLPPENEIDFVPESEDRDKTIALINNSGSGTLPKGVALPHRTACVR
FSHARDPIFGNQIIPDTAILSVVFFHHGFMFTTLGYLIGCFRVVLMYRPEEELFLRS
LDYQKISALLVPTLFSFAKSTLIDKYLNSLHIEASGGAPLSKEVEGAVAKRPHLP
RGQMISGVYNNPEATNALIDKDWLHSGDIAVWDEHFFIVDRLKSLLIKYGYQVA
PARESLILQHPNIPDAGVAGLPDDAGELPAAVVLEHGTWTEKEIVDIVVASQVTT
AKLRGGVVFVDEVPKGLTGKLDARKIRELIIKAKGGKSKL"

misc_feature
3359..3464
/notes="Heliothis virescens p63 chaperonin gene double
poly-adenylation sequence"
4381..5241
/codon_start=1
/product="beta-lactamase"
/protein_id="AAL30783.1"
/db_xref="GI:16904160"

rep_origin
5529..5948
/note="ColE1"

ORIGIN

Query Match 81.3%; Score 574.6; DB 12; Length 6182;
Best Local Similarity 99.3%; Pred. No. 1.8e-145;
Matches 577; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 45 ATCTAGCTGATCAGGATCATATCGTCGGGCTCTTTTCCGGCTCAGTCATCGCCCAAGC 104
|||
Db 3721 ATCCAGCTGATCAGGATCATATCGTCGGGCTCTTTTCCGGCTCAGTCATCGCCCAAGC 3780
|||
Qy 105 TGGCGCTATCTGGGCATCGGGAGGAAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 164
|||
Db 3781 TGGCGCTATCTGGGCATCGGGAGGAAAGCCCGTGCCTTTTCCCGCGAGGTTGAAGCG 3840
|||
Qy 165 GCATGGAAGAGTTTCCGAGGATGACTGCTGTGCATGACGTTGAGCGAAACGCAACG 224
|||
Db 3841 GCATGGAAGAGTTTCCGAGGATGACTGCTGTGCATGACGTTGAGCGAAACGCAACG 3900
|||
Qy 225 TTTACCATGATGATTCGGGAGGATGAGCCATCGACGCTTTTAAACGTTGACTTTCGTT 284
|||
Db 3901 TTTACCATGATGATTCGGGAGGATGAGCCATCGACGCTTTTAAACGTTGACTTTCGTT 3960
|||
Qy 285 CAGGCCACCTGGGATACAGTTCGTCGGGCTTTTCCGGACACAGTTCGGATGGTGTCAGC 344
|||
Db 3961 CAGGCCACCTGGGATACAGTTCGTCGGGCTTTTCCGGACACAGTTCGGATGGTGTCAGC 4020
|||

```
Qy 345 CCGAAGCGCATCAGCAACCCGAAACAATACCGCGCGACAGCCGGAACCTGCGGTGCGGTGTG 404
Dy 4021 CCGAAGCGCATCAGCAACCCGAAACAATACCGCGCGACAGCCGGAACCTGCGGTGCGGTGTG 4080
Qy 405 CAGATTAATGACAGCGGTGCGCGCTGGGATATTACGTACGCGAGGACGGGTATCTCTGGC 464
Dy 4081 CAGATTAATGACAGCGGTGCGCGCTGGGATATTACGTACGCGAGGACGGGTATCTCTGGC 4140
Qy 465 TGGATGCCGAGAAATGACATGATACCCCGTGAGTTTACCCGCGCGCGCGCTCGTTC 524
Dy 4141 TGGATGCCGAGAAATGACATGATACCCCGTGAGTTTACCCGCGCGCGCGCTCGTTC 4200
Qy 525 ATTACAGCTTTTGAACCCCGTGAGGACCGGACAGCTCGCGGTGCAAAATGTGTTTACAGC 584
Dy 4201 ATTACAGCTTTTGAACCCCGTGAGGACCGGACAGCTCGCGGTGCAAAATGTGTTTACAGC 4260
Qy 585 GTGATGGAGCAGATGAAGATGCTGCACACGCTGCAGAACAC 625
Dy 4261 GTGATGGAGCAGATGAAGATGCTGCACACGCTGCATGAAC 4301

RESULT 10
AF434926 AF434926 6182 bp DNA circular SYN 04-APR-2002
LOCUS Expression vector 409-MUT, complete sequence.
DEFINITION AF434926
ACCESSION AF434926
VERSION AF434926.1 GI:16904161
KEYWORDS
SOURCE Expression vector 409-MUT
ORGANISM Expression vector 409-MUT
REFERENCE Expression vector 409-MUT
REFERENCE 1 (bases 1 to 6182)
AUTHORS other sequences; artificial sequences; vectors.
TITLE Zieler, H. and Huynh, C.O.
intron-dependent stimulation of marker gene expression in cultured
insect cells
JOURNAL Insect Mol. Biol. 11 (1), 87-95 (2002)
MEDLINE 21830665
PUBMED 11841506
REFERENCE 2 (bases 1 to 6182)
AUTHORS Zieler, H. and Huynh, C.O.
TITLE Direct Submission
JOURNAL Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of
Parasitic Diseases, National Institutes of Health, 4 Center Drive
MSC 0425, Bethesda, MD 20892-0425, USA
FEATURES
Location/Qualifiers
source
1. 6182
/organism="Expression vector 409-MUT"
/mol_type="genomic DNA"
/db_xref="taxon:176968"
/feature="insect-specific"
97. 614
/feature="insect-specific"
enhancer
1. 614
/feature="Autographa californica nuclear polyhedrosis virus
hr5 enhancer"
615. 1201
/feature="Autographa californica nuclear polyhedrosis virus
immediate early promoter"
1266. 1398
/feature="artificial intron with mutated 3' splice site"
1541. 3193
/codon_start=1
/product="firefly luciferase"
/protein_id="AAL30784.1"
/db_xref="GI:16904162"
/translation="MEDAKNIKKGPAPPYPLEDGTAGCQLHKKRYALVPGTIAFTD
AHIEVNTIYBYFMSVRLAEMKRYGLNTHRIIVVCSNSLQFPMVLGALFVGAV
APANDIYNERELNSMNIQPTVVFVSKGLQKLNQKLP11OKIILIMDSKTDYQ
PQSMYFTPLPFGNIDYDFVPSFDRKDTIALIMNSSGSTGLPKGVALPHRTACVR
FSHARDPIFGNIIIPDNTAIVSVPHFGFMFTLGLYICGFRVLMVRFEEELFLRS
LDQVKIQSALLVPTLFSFPAKSTLIDKVLNLHEIASGGAPLSKEVCEAVAKRFLPL
GIRQYGLTETTSAILITPEGDKPGAVGVVPEAKVVDLDTGKTLGVNORGLCV
RGPIMMSYNNPENATNALIDKGLHSGDIAYNDEDEHFFIVDRLSLILYKGYQVA
PAELSELILQHPNIFDAGVAGLPDDAGELPAAVVVLHGKMTKEIVDVVASQVTT
AKLRGGVVVFDEVPKGLTGKLDARKIREILIKAKGGKSKL"
```

```
misc_feature 3359..3464
/note="Heliothis virescens p63 chaperonin gene double
poly-adenylation sequence"
CDS
4381..5241
/codon_start=1
/product="beta-lactamase"
/protein_id="AAL30785.1"
/db_xref="GI:16904163"
/translation="MSIQHPRVALIPFAAFCLVPFAHPETLVKVKDAEDQLGARVGY
IEDLNSGKILSEPRPEEPPMSTFKVLLCGAVLSRI DAGQEQLGRIHYSONDLVE
YSPVTEKHLTDGNTVRELCSAALTMSDNTAANLLLTITIGSPKELTAFILNMGDHVTSL
DRWPELNEAIPNDERDITMPVAMATLRLKLTGLLELTLASRQLGLDWMDEADKRVAGPL
LRSLALPAGWFPIADKSGAGERGSRGIIAALGPDGKPSRIIVVIYTTGSSQATMDERNRQIA
EIGASLIKHW"
rep_origin 5529..5948
/note="ColE1"
ORIGIN
Query Match 81.3%; Score 574.6; DB 12; Length 6182;
Best Local Similarity 99.3%; Pred. No. 1.8e-145;
Matches 577; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 45 ATCTAGCTGCATCGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 104
Dy 3721 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 3780
Qy 105 TGGCGCTATCTGGGCATCGGGGAGGAAGCCCGTGCCTTTTCCCGGAGGTTGAAGCG 164
Dy 3781 TGGCGCTATCTGGGCATCGGGGAGGAAGCCCGTGCCTTTTCCCGGAGGTTGAAGCG 3840
Qy 165 GCATGGAAGAGTTTGGCGAGGATGACTGCTGCTGCAATTGACGTTGAGCGGAAACGCAAG 224
Dy 3841 GCATGGAAGAGTTTGGCGAGGATGACTGCTGCTGCAATTGACGTTGAGCGGAAACGCAAG 3900
Qy 225 TTACCATGATGATTCCGGGAAGGTGGCCATGACGCCCTTTAACGCTGACTGTTGTT 284
Dy 3901 TTACCATGATGATTCCGGGAAGGTGGCCATGACGCCCTTTAACGCTGACTGTTGTT 3960
Qy 285 CAGGCCACCTGGGATACAGTTTCGTCGCGCTTTCCGGACACAGTTTCCGGATCGTCAGC 344
Dy 3961 CAGGCCACCTGGGATACAGTTTCGTCGCGCTTTCCGGACACAGTTTCCGGATCGTCAGC 4020
Qy 345 CCGAAGCGCATCAGCAACCCGAAACAATACCGCGCGACAGCCGGAACCTGCGGTGCGGTG 404
Dy 4021 CCGAAGCGCATCAGCAACCCGAAACAATACCGCGCGACAGCCGGAACCTGCGGTGCGGTG 4080
Qy 405 CAGATTAATGACAGCGGTGCGCGCTGGGATATTACGTACGCGAGGACGGGTATCTCTGGC 464
Dy 4081 CAGATTAATGACAGCGGTGCGCGCTGGGATATTACGTACGCGAGGACGGGTATCTCTGGC 4140
Qy 465 TGGATGCCGAGAAATGACATGATACCCCGTGAGTTTACCCGCGCGCGCTCGTTC 524
Dy 4141 TGGATGCCGAGAAATGACATGATACCCCGTGAGTTTACCCGCGCGCGCTCGTTC 4200
Qy 525 ATTACAGCTTTTGAACCCCGTGAGGACCGGACAGCTCGCGGTGCAAAATGTGTTTACAGC 584
Dy 4201 ATTACAGCTTTTGAACCCCGTGAGGACCGGACAGCTCGCGGTGCAAAATGTGTTTACAGC 4260
Qy 585 GTGATGGAGCAGATGAAGATGCTGCACACGCTGCAGAACAC 625
Dy 4261 GTGATGGAGCAGATGAAGATGCTGCACACGCTGCATGAAC 4301

RESULT 11
AF434927 AF434927 6182 bp DNA circular SYN 04-APR-2002
LOCUS Expression vector 410-FOR, complete sequence.
DEFINITION AF434927
ACCESSION AF434927
VERSION AF434927.1 GI:16904164
KEYWORDS
SOURCE Expression vector 410-FOR
ORGANISM Expression vector 410-FOR
other sequences; artificial sequences; vectors.
```

```
REFERENCE 1 (bases 1 to 6182)
AUTHORS Zieler,H. and Huynh,C.Q.
TITLE Intron-dependent stimulation of marker gene expression in cultured insect cells
JOURNAL Insect Mol. Biol. 11 (1), 87-95 (2002)
MEDLINE 21830665
PUBMED 11841506
REFERENCE 2 (bases 1 to 6182)
AUTHORS Zieler,H. and Huynh,C.Q.
TITLE Direct Submission
JOURNAL Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of Parasitic Diseases, National Institutes of Health, 4 Center Drive MSC 0425, Bethesda, MD 20892-0425, USA
FEATURES
    source
        1. .6182
            /organism="Expression vector 410-FOR"
            /mol_type="genomic DNA"
            /db_xref="taxon:176969"
            /notes="insect-specific"
            97. .614
            /note="Autographa californica nuclear polyhedrosis virus hr-5 enhancer"
        615. .1201
            /note="Autographa californica nuclear polyhedrosis virus immediate early promoter"
            1328. .2980
            /codon_start=1
            /product="firefly luciferase"
            /protein_id="AAL30786.1"
            /db_xref="GI:16904165"
            translation="MEDAKNIKGPAPFYPLEDGTAGEQLHKMKRYALVPGTIAPTD
            AHIEVNTIAYEPMFVRVLAEMKRYGLNTHRIVVCSNSLQFPFVLGALFIVAV
            APANDYINRELINSMNI SQPTVVFVSKGLQKILNVOKKLPITQKIIIMDSKTDYQG
            FOSMYTFVSHLPPGNEVDVFPESDRDKTIALIMNSGSTGLPGVALPHRTACVR
            FSHARDPIQNGOIIPTAILSVVPHHGGMTTGLYLCGRVVLWYFERELFLRS
            LDQKJQSALLPTLSPFAKSTLIDKYLNSLHEIASGAPLSKVGVAARFHLPL
            GRQGGVLTLSAILITPEGGDKPGAVGVVFFFAKVVDDLTGKTLGVNGRGEUCV
            RQPMINSVYNNPEINATLIDKQWLSHGDIAWDEDEHFFIVDRLSKIRYKGYQVA
            PAELSIILQHNPNI FADGABLPDDDABLPAAVVLHGKTMTEKEIVDYVASQVTT
            AKKLRGVVFDVBPVKGLTKLDARKIRILIKAKKGGKSL"
            3196. .3328
            /note="artificial intron"
        3359. .3464
            /note="Heliothis virescens p63 chaperonin gene double
            poly-adenylation sequence"
            4381. .5241
            /codon_start=1
            /product="beta-lactamase"
            /protein_id="AAL30787.1"
            /db_xref="GI:16904166"
            translation="MSIQHFRVALIPFFAAFLPVPFAHPETLVKVKDAEDQLGARVGY
            IELDLSGKILSFEPPEFPPMSTFKLLCGAVLSRI DAGOEOLGRRIRHYSDNLVLE
            YSPVTEKHTLDGKTVRELCSAAITMSDNTAANLLTIGPKELTAFLHMGDHYVREL
            DRWEPELNAIPNDERDTPVMATTLKRLTGELLTGLASQQLIDWAEADKVAQPL
            LRSLALPAGFIADKSGRSGRGIILALGPDGPKPRIVVIYTTGSGQTMDERNRQIA
            ETGASLIKKHW"
            5529. .5948
            /note="ColE1"
    rep_origin
        81.3%; Score 574.6; DB 12; Length 6182;
        Best Local Similarity 99.3%; Pred. No. 1.8e-145;
        Matches 577; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
    ORIGIN
        45 ATCTAGCTGATCAGGATCATATCGTCGGGCTCTTTTTCGGGCTCAGTCATCGCCCAAGC 104
        |||
        3721 ATCCAGCTGATCAGGATCATATCGTCGGGCTCTTTTTCGGGCTCAGTCATCGCCCAAGC 3780
        |||
        105 TGGCGCTATCTGGGCATCGGGGAGGAAGACCGCGCTCTTTTTCGGGCTGAGTTGAAGCG 164
        |||
        3781 TGGCGCTATCTGGGCATCGGGGAGGAAGACCGCGCTCTTTTTCGGGCTGAGTTGAAGCG 3840
        |||
        165 GCATGGAAGAGTTTGGCCGAGGATGACTGCTGCTGATTCGTTGACGTTGAGCGAAAAACGCAGC 224
```

```
Db 3841 GCATGGAAGAGTTTGGCCGAGGATGACTGCTGCTGATTCGTTGAGCGAAAAACGCAGC 3900
Qy 225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGTCAGCCCTTTTAAACGGTGAACCTGTTGTT 284
|||
Db 3901 TTTACCATGATGATTCGGGAAGGTGTGGCCATGTCAGCCCTTTTAAACGGTGAACCTGTTGTT 3960
Qy 285 CAGGCCACCTGGGATACCAAGTTTCGTCGGGCTTTTCCGACACAGTTCGGATGTCAGC 344
|||
Db 3961 CAGGCCACCTGGGATACCAAGTTTCGTCGGGCTTTTCCGACACAGTTCGGATGTCAGC 4020
Qy 345 CCGAAGCGATCAGCAACCCGAAACAATACCGGCGACAGCGGAACCTCGCTGCCGGTGTG 404
|||
Db 4021 CCGAAGCGATCAGCAACCCGAAACAATACCGGCGACAGCGGAACCTCGCTGCCGGTGTG 4080
Qy 405 CAGATTAAATGACACGGGTGCGGCTGGGATATTACGTGACGAGGACCGGGTATCTCTGGC 464
|||
Db 4081 CAGATTAAATGACACGGGTGCGGCTGGGATATTACGTGACGAGGACCGGGTATCTCTGGC 4140
Qy 465 TGGATGCCCGAGAATGACATGATACCCCGTGTAGTTACCGCGCGCGCGCTCGTTC 524
|||
Db 4141 TGGATGCCCGAGAATGACATGATACCCCGTGTAGTTACCGCGCGCGCGCTCGTTC 4200
Qy 525 ATTACGTTTTCGACCCCGTGGAGGACGGGACAGCTCGCGGTGCAATGTGTTTACAGC 584
|||
Db 4201 ATTACGTTTTCGACCCCGTGGAGGACGGGACAGCTCGCGGTGCAATGTGTTTACAGC 4260
Qy 585 GTGATGGACGACATGAAGATGCTCGACACGCTGCAGAACAC 625
|||
Db 4261 GTGATGGACGACATGAAGATGCTCGACACGCTGCAGAACAC 4301
|||
RESULT 12
AF434931 6182 bp DNA circular SYN 04-APR-2002
LOCUS Expression vector 409-REV, complete sequence.
DEFINITION AF434931
ACCESSION AF434931 GI:16904176
VERSION AF434931.1
KEYWORDS Expression vector 409-REV
SOURCE Expression vector 409-REV
ORGANISM Expression vector 409-REV
other sequences; artificial sequences; vectors.
REFERENCE 1 (bases 1 to 6182)
AUTHORS Zieler,H. and Huynh,C.Q.
TITLE Intron-dependent stimulation of marker gene expression in cultured insect cells
JOURNAL Insect Mol. Biol. 11 (1), 87-95 (2002)
MEDLINE 21830665
PUBMED 11841506
REFERENCE 2 (bases 1 to 6182)
AUTHORS Zieler,H. and Huynh,C.Q.
TITLE Direct Submission
JOURNAL Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of Parasitic Diseases, National Institutes of Health, 4 Center Drive MSC 0425, Bethesda, MD 20892-0425, USA
FEATURES
    source
        1. .6182
            /organism="Expression vector 409-REV"
            /mol_type="genomic DNA"
            /db_xref="taxon:176973"
            /notes="insect-specific"
            97. .614
            /note="Autographa californica nuclear polyhedrosis virus hr-5 enhancer"
        615. .1201
            /note="Autographa californica nuclear polyhedrosis virus immediate early promoter"
            complement(1236..1368)
            /notes="artificial intron"
            1541. .3193
            /codon_start=1
            /product="firefly luciferase"
            /protein_id="AAL30794.1"
```



```
/db_xref="GI:16904177"
/translation="MEDAKNIKKGPAPFPYLEDGTAGBQLHKAHKRYALVPGTTIAFTD
AHIVNTIYASYFMSVLAEMAKRYGLNTNHRIVVCSNSLOQFMPVLGALFIGVAV
APANDIYNERLLSNMISQPTVVFVSKGLOKILNVQKLPILQKIIIMDSKTDYQG
FQSMYTFVTHLPGFNEYDFVPSFDRKDTIALIMNSGSGTLPKGVLPHTACVR
FSHARDPIFGNQIIPDTAILSVVPHFGFMFTTGLYICGFRVLMVRFEEELFLRS
LDQYIKSALAPVTLTFQFFAKSTIIDKYLNSLHEIASGGAPLSKEVGEAVAKRFLP
GIQGYGLITETTSAILITPEGDKPGAVKVPVFEAKVVDLDTGKTLGVNORGLCV
RGPIMSGYVNNPNTALIDKDWLHSGDIAWDEDEHFFIVDRLSKSLIKYKGQYVA
PAELISILLOHPNIFDAGVAGLPDDAGELPAAVVVLEHGKTMTEKEIVDVASQVTT
AKKLRGGVVFVDEVPKGLTGKLDARKIREILIKAKGGKSKL"
misc_feature
3359..3464
/note="Heliothis virescens p63 chaperonin gene double
poly-adenylation sequence"
CDS
4381..5241
/codon_start=1
/product="beta-lactamase"
/protein_id="AAL30795.1"
/db_xref="GI:16904178"
/translation="MSIQHFRVALIPFFAARCLPVFAHPETLVKVKDAEDQLGARVGY
IELDLSGKILSFRRPEERPMSTMTEFVKLLCGAVLSRIDAGQEOLGRRIHYSONDLVE
YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLTTIGGKELTAFLHNMGDHVTSL
DRWPELNEALPNDERDTTPVAMATTLIKLLTGELLTLASRQQLIDWMEADKVAAGL
LRSLAPAGWFIADKSGAGERSGRIIAALGPDGKPSRIIVYITGSGQATMDERNRQIA
EIGASLTKHW"
rep_origin
5529..5948
/note="ColE1"
```

ORIGIN

```
Query Match      81.3%; Score 574.6; DB 12; Length 6182;
Best Local Similarity 99.3%; Pred. No. 1.8e-145;
Matches 577; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 45 ATCTAGCTGCATCAGGATCATATCGTCGGGCTTTTTCGGCTCAGTCATCGCCCAAGC 104
Db |||
Qy 3721 ATCCAGCTGCATCAGGATCATATCGTCGGGCTTTTTCGGCTCAGTCATCGCCCAAGC 3780
Db |||
Qy 105 TGGCGCTATCTGGGCATCGGGAGGAAGAACCCGTCGCTTTTCGGCGAGGTTGAAGCG 164
Db |||
Qy 3781 TGGCGCTATCTGGGCATCGGGAGGAAGAACCCGTCGCTTTTCGGCGAGGTTGAAGCG 3840
Db |||
Qy 165 GCATGGAAGAGTTTGGCAGGATGACCTGCTGTCATGACGTTGACGAGGAAACGACG 224
Db |||
Qy 3841 GCATGGAAGAGTTTGGCAGGATGACCTGCTGTCATGACGTTGACGAGGAAACGACG 3900
Db |||
Qy 225 TTTACCATGATGATTTCGGGAAGGTGTGCCATGCACGCTTTTAAACGGTGAACTGTCGTT 284
Db |||
Qy 3901 TTTACCATGATGATTTCGGGAAGGTGTGCCATGCACGCTTTTAAACGGTGAACTGTCGTT 3960
Db |||
Qy 285 CAGGCCACCTGGGATACCAAGTTCGTCGGGCTTTTCGGGACACAGTTCGGGATGTCGACG 344
Db |||
Qy 3961 CAGGCCACCTGGGATACCAAGTTCGTCGGGCTTTTCGGGACACAGTTCGGGATGTCGACG 4020
Db |||
Qy 345 CCGAAGCGCATCAGCAACCCGAAACAATACCGCGGACAGCCGGAACCTGCCGTGCGGTG 404
Db |||
Qy 4021 CCGAAGCGCATCAGCAACCCGAAACAATACCGCGGACAGCCGGAACCTGCCGTGCGGTG 4080
Db |||
Qy 405 CAGATTATGACAGCGTTCGGCGCTGGGATATTACGTCAGGAGGCGGTATCTCTGGC 464
Db |||
Qy 4081 CAGATTATGACAGCGTTCGGCGCTGGGATATTACGTCAGGAGGCGGTATCTCTGGC 4140
Db |||
Qy 465 TGGATGCGCGAGAAATGACATGGATACCCCGTGAAGTTACCCGGCGGCGCGCTCGTTC 524
Db |||
Qy 4141 TGGATGCGCGAGAAATGACATGGATACCCCGTGAAGTTACCCGGCGGCGCGCTCGTTC 4200
Db |||
Qy 525 ATTACGCTTTTGAACCCCGTGGAGAGACCGGCGAGACTCGCGGTGCNAATGTGTTTACAGC 584
Db |||
Qy 4201 ATTACGCTTTTGAACCCCGTGGAGAGACCGGCGAGACTCGCGGTGCNAATGTGTTTACAGC 4260
Db |||
Qy 585 GTGATGAGCAGATGAAGATGCTGCACACGCTGCAGACAC 625
Db |||
Qy 4261 GTGATGAGCAGATGAAGATGCTGCACACGCTGCATGAAAC 4301
Db |||
```

```
RESULT 13
AP434932
LOCUS AP434932 6182 bp DNA circular SYN 04-APR-2002
DEFINITION Expression vector 410-REV, complete sequence.
ACCESSION AP434932
VERSION AP434932.1 GI:16904179
KEYWORDS
SOURCE
ORGANISM
Expression vector 410-REV
Other sequences; artificial sequences; vectors.
REFERENCE
1 (bases 1 to 6182)
Zieler,H. and Huynh,C.Q.
TITLE Intron-dependent stimulation of marker gene expression in cultured insect cells
JOURNAL Insect Mol. Biol. 11 (1), 87-95 (2002)
MEDLINE 21830665
PUBMED 11841506
REFERENCE
2 (bases 1 to 6182)
Zieler,H. and Huynh,C.Q.
TITLE Direct Submision
JOURNAL Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of Parasitic Diseases, National Institutes of Health, 4 Center Drive MSC 0425, Bethesda, MD 20892-0425, USA
FEATURES
source
1..6182
/organism="Expression vector 410-REV"
/mol_type="genomic DNA"
/db_xref="taxon:176974"
/notes="insect-specific"
97..614
/note="Autographa californica nuclear polyhedrosis virus hr5 enhancer"
615..1201
/notes="Autographa californica nuclear polyhedrosis virus immediate early promoter"
1328..2980
/codon_start=1
/product="firefly luciferase"
/protein_id="AAL30796.1"
/db_xref="GI:16904180"
/translation="MEDAKNIKKGPAPFPYLEDGTAGBQLHKAHKRYALVPGTTIAFTD
AHIVNTIYASYFMSVLAEMAKRYGLNTNHRIVVCSNSLOQFMPVLGALFIGVAV
APANDIYNERLLSNMISQPTVVFVSKGLOKILNVQKLPILQKIIIMDSKTDYQG
FQSMYTFVTHLPGFNEYDFVPSFDRKDTIALIMNSGSGTLPKGVLPHTACVR
FSHARDPIFGNQIIPDTAILSVVPHFGFMFTTGLYICGFRVLMVRFEEELFLRS
LDQYIKSALAPVTLTFQFFAKSTIIDKYLNSLHEIASGGAPLSKEVGEAVAKRFLP
GIQGYGLITETTSAILITPEGDKPGAVKVPVFEAKVVDLDTGKTLGVNORGLCV
RGPIMSGYVNNPNTALIDKDWLHSGDIAWDEDEHFFIVDRLSKSLIKYKGQYVA
PAELISILLOHPNIFDAGVAGLPDDAGELPAAVVVLEHGKTMTEKEIVDVASQVTT
AKKLRGGVVFVDEVPKGLTGKLDARKIREILIKAKGGKSKL"
complement(3166..3298)
/notes="artificial intron"
3359..3464
/note="Heliothis virescens p63 chaperonin gene double
poly-adenylation sequence"
4381..5241
/codon_start=1
/product="beta-lactamase"
/protein_id="AAL30797.1"
/db_xref="GI:16904181"
/translation="MSIQHFRVALIPFFAARCLPVFAHPETLVKVKDAEDQLGARVGY
IELDLSGKILSFRRPEERPMSTMTEFVKLLCGAVLSRIDAGQEOLGRRIHYSONDLVE
YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLTTIGGKELTAFLHNMGDHVTSL
DRWPELNEALPNDERDTTPVAMATTLIKLLTGELLTLASRQQLIDWMEADKVAAGL
LRSLAPAGWFIADKSGAGERSGRIIAALGPDGKPSRIIVYITGSGQATMDERNRQIA
EIGASLTKHW"
5529..5948
/note="ColE1"
rep_origin
ORIGIN
Query Match      81.3%; Score 574.6; DB 12; Length 6182;
Best Local Similarity 99.3%; Pred. No. 1.8e-145;
```


Matches	577;	Conservative	0;	Mismatches	4;	Indels	0;	Gaps	0;
Qy	45	ATCTAGCTGCATCAGGATCATATCGTCGGGCTCTTTTTCGGGCTCAGTCAATCGCCCAAGC	104						
Db	3721	ATCCAGCTGCATCAGGATCATATCGTCGGGCTCTTTTTCGGGCTCAGTCAATCGCCCAAGC	3780						
Qy	105	TGGCGCTATCTGGGCATTCGGGAGGAGAAAGCCGCTCTTTTTCGGGAGGTTGAAGCG	164						
Db	3781	TGGCGCTATCTGGGCATTCGGGAGGAGAAAGCCGCTCTTTTTCGGGAGGTTGAAGCG	3840						
Qy	165	GCATGGAACAGTTTCGGGAGGATGACTGCTGCTGATTAACCGTTCGAGCGAAACCGCAG	224						
Db	3841	GCATGGAACAGTTTCGGGAGGATGACTGCTGCTGATTAACCGTTCGAGCGAAACCGCAG	3900						
Qy	225	TTTACCATGATGATTCGGGAGGATGCTGCCATGCAGCCCTTTTAAACCGTGAACCTGTCGTT	284						
Db	3901	TTTACCATGATGATTCGGGAGGATGCTGCCATGCAGCCCTTTTAAACCGTGAACCTGTCGTT	3960						
Qy	285	CAGGCCACTTGGGATACCATGTCGTCGGGCTTTTTCGGGACACAGTTTCGGATGGTCAGC	344						
Db	3961	CAGGCCACTTGGGATACCATGTCGTCGGGCTTTTTCGGGACACAGTTTCGGATGGTCAGC	4020						
Qy	345	CCGAAGCGCATCAGCAACCGAACATACCGCGGACAGCCGGAACCTGCCGTGCGGTTG	404						
Db	4021	CCGAAGCGCATCAGCAACCGAACATACCGCGGACAGCCGGAACCTGCCGTGCGGTTG	4080						
Qy	405	CAGATTAAATGACAGCGGTGCGGCTGGGATATTAACGTCAGCAGAGCGGATCTCTGGC	464						
Db	4081	CAGATTAAATGACAGCGGTGCGGCTGGGATATTAACGTCAGCAGAGCGGATCTCTGGC	4140						
Qy	465	TGGATGCCGAGAAATGGACATGATACCCCGTGAGTTACCCGCGGCGCGCTCGTTC	524						
Db	4141	TGGATGCCGAGAAATGGACATGATACCCCGTGAGTTACCCGCGGCGCGCTCGTTC	4200						
Qy	525	ATTACGCTTTTGAACCCGTGGAGGAGCGGACACTCGCGTGCAATGTGTTTACAGC	584						
Db	4201	ATTACGCTTTTGAACCCGTGGAGGAGCGGACACTCGCGTGCAATGTGTTTACAGC	4260						
Qy	585	GTGATGAGCAGATGAAGATGCTGCACACGCTGCAGAACAC	625						
Db	4261	GTGATGAGCAGATGAAGATGCTGCACACGCTGCAGAACAC	4301						

RESULT 14
AF434924
LOCUS AF434924 7421 bp DNA circular SYN 04-APR-2002
DEFINITION Expression vector pACTIN-LUC, complete sequence.
ACCESSION AF434924
VERSION AF434924.1 GI:16904155
KEYWORDS Expression vector pACTIN-LUC
SOURCE Expression vector pACTIN-LUC
ORGANISM Expression vector pACTIN-LUC
REFERENCE 1 (bases 1 to 7421)
Zieler, H. and Huynh, C.Q.
TITLE Intron-dependent stimulation of marker gene expression in cultured insect cells
JOURNAL Insect Mol. Biol. 11 (1), 87-95 (2002)
MEDLINE 21830665
PUBMED 11841506
REFERENCE 2 (bases 1 to 7421)
Zieler, H. and Huynh, C.Q.
TITLE Direct Submision
JOURNAL Submitted (17-OCT-2001) Medical Entomology Section, Laboratory of Parasitic Diseases, National Institutes of Health, 4 Center Drive MSC 0425, Bethesda, MD 20892-0425, USA
FEATURES
source
1. .7421
/organism="Expression vector pACTIN-LUC"
/mol_type="genomic DNA"
/db_xref="taxon:176966"
/notes="insect-specific"
103. .2660
promoter

CDS	misc_feature	CDS	rep_origin	ORIGIN	Query Match	Best Local Similarity	Matches	Conservative	0;	Mismatches	4;	Indels	0;	Gaps	0;
/note="Drosophila melanogaster actin5C promoter"	2780. .4432	/codon_start=1	5620. .6480	6768. .7187	81.3%;	Score 574.6;	DB 12;	Length 7421;							
/product="firefly luciferase"	/protein_id="AAL30780.1"	/db_xref="GI:16904156"	/translaction="MSIQHFRVALIPFFAFCPLPVFAHPETLVKVKDAEDQLGARVGY	/protein_id="AAL30781.1"	99.3%;	Pred. No. 1.8e-145;									
/db_xref="GI:16904156"	/translaction="MSIQHFRVALIPFFAFCPLPVFAHPETLVKVKDAEDQLGARVGY	/protein_id="AAL30781.1"	/db_xref="GI:16904157"	/translaction="MSIQHFRVALIPFFAFCPLPVFAHPETLVKVKDAEDQLGARVGY	577;	Conservative	0;	Mismatches	4;	Indels	0;	Gaps	0;		
45	ATCTAGCTGCATCAGGATCATATCGTCGGGCTCTTTTTCGGGCTCAGTCAATCGCCCAAGC	104			4960	ATCCAGCTGCATCAGGATCATATCGTCGGGCTCTTTTTCGGGCTCAGTCAATCGCCCAAGC	5019								
105	TGGCGCTATCTGGGCATTCGGGAGGAGAAAGCCGCTCTTTTTCGGGAGGTTGAAGCG	164			5020	TGGCGCTATCTGGGCATTCGGGAGGAGAAAGCCGCTCTTTTTCGGGAGGTTGAAGCG	5079								
165	GCATGGAACAGTTTCGGGAGGATGACTGCTGCTGATTAACCGTTCGAGCGAAACCGCAG	224			5080	GCATGGAACAGTTTCGGGAGGATGACTGCTGCTGATTAACCGTTCGAGCGAAACCGCAG	5139								
225	TTTACCATGATGATTCGGGAGGATGCTGCCATGCAGCCCTTTTAAACCGTGAACCTGTCGTT	284			5140	TTTACCATGATGATTCGGGAGGATGCTGCCATGCAGCCCTTTTAAACCGTGAACCTGTCGTT	5199								
285	CAGGCCACTTGGGATACCATGTCGTCGGGCTTTTTCGGGACACAGTTTCGGATGGTCAGC	344			5200	CAGGCCACTTGGGATACCATGTCGTCGGGCTTTTTCGGGACACAGTTTCGGATGGTCAGC	5259								
345	CCGAAGCGCATCAGCAACCGGAACATACCGCGGACAGCCGGAACCTGCCGTGCGGTTG	404			5260	CCGAAGCGCATCAGCAACCGGAACATACCGCGGACAGCCGGAACCTGCCGTGCGGTTG	5319								
405	CAGATTAAATGACAGCGGTGCGGCTGGGATATTAACGTCAGCAGAGCGGATCTCTGGC	464			5320	CAGATTAAATGACAGCGGTGCGGCTGGGATATTAACGTCAGCAGAGCGGATCTCTGGC	5379								
465	TGGATGCCGAGAAATGGACATGATACCCCGTGAGTTACCCGCGGCGCGCTCGTTC	524			5380	TGGATGCCGAGAAATGGACATGATACCCCGTGAGTTACCCGCGGCGCGCTCGTTC	5439								
525	ATTACGCTTTTGAACCCGTGGAGGAGCGGACACTCGCGTGCAATGTGTTTACAGC	584			5440	ATTACGCTTTTGAACCCGTGGAGGAGCGGACACTCGCGTGCAATGTGTTTACAGC	5499								

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 12, 2005, 13:46:55 ; Search time 3474 Seconds
(without alignments)
1204.737 Million cell updates/sec

Title: US-10-826-523-40

Perfect score: 707

Sequence: 1 ggatccatgcgcaatttt.....aaaattgacgatgggatcc 707

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	707	100.0	707	10	ACA55353 Eukaryoti
2	707	100.0	4613	10	ACA55360 Transform
C 3	705.4	99.8	8999	10	ACA55361 Transform
C 4	705.4	99.8	9009	10	ACA55362 Transform
C 5	705.4	99.8	9012	10	ACA55363 Transform
6	685	96.9	4943	10	ACA55369 Transform
7	685	96.9	4943	10	ACA55367 Transform
8	681.8	96.4	4944	10	ACA55368 Transform
9	585.4	82.1	48502	12	ADP70043 Nucleic a
10	580.6	82.1	4164	4	AAS44733 Bacteriop
C 11	566.2	80.1	26565	6	ABR78924 E. coli C
C 12	566.2	80.1	26565	10	ADH80491 Enterohae
13	566.2	80.1	46819	9	ACD19059 E. coli 0
14	564.6	79.9	46897	10	ADC00585 Enterohae
C 15	561.4	79.4	22306	9	ACD19238 E. coli 0
C 16	561.4	79.4	91740	10	ADC00956 Enterohae
C 17	519.2	73.4	3661	10	ACA55354 Transform
C 18	511	72.3	4941	10	ACA55366 Transform
C 19	511	72.3	4952	10	ACA55365 Transform
C 20	499	70.6	4951	10	ACA55364 Transform

21	425.8	60.2	1427	6	ABQ50926	Abq50926 Oligonuc1
C 22	425.8	60.2	1427	6	ABQ50927	Abq50927 Oligonuc1
23	392.6	55.5	3822	5	AAS94485	Aas94485 DNA encod
C 24	382.6	54.1	1427	6	ABQ50928	Abq50928 Oligonuc1
25	382.6	54.1	1427	6	ABQ50929	Abq50929 Oligonuc1
C 26	223	31.5	34063	9	ACD19083	Adc19083 E. coli 0
C 27	223	31.5	44029	10	ADC00710	Adc00710 Enterohae
28	223	31.5	49650	10	ADC00365	Adc00365 Enterohae
C 29	222.2	31.4	134141	6	ABN83487	Abn83487 Escherich
C 30	171.2	24.2	822	5	AAS80174	Aas80174 DNA encod
C 31	96	13.6	847	2	Aaz27249	Aaz27249 Human sec
C 32	78	11.0	9980	10	ACA55352	ACA55352 pIAO-P/L
C 33	73.8	10.4	5194	4	AAC85595	Aac85595 Plasmid p
C 34	73.8	10.4	5194	9	ADA09845	Ada09845 PiggyBac
C 35	73.4	10.4	5194	4	AAC85596	Aac85596 Plasmid p
C 36	73.4	10.4	5194	9	ADA09846	Ada09846 PiggyBac
C 37	72.8	10.3	7670	10	ACA55351	ACA55351 pIAO-P/L
C 38	67.8	9.6	2480	8	ABV76195	Abv76195 T-richoplu
C 39	67.8	9.6	5679	4	AAC85602	Aac85602 Helper pl
C 40	67.8	9.6	5679	9	ADA09857	Ada09857 Helper pl
C 41	67.8	9.6	6723	4	AAC85593	Aac85593 psep70/op
C 42	67.8	9.6	6723	4	AAC85592	Aac85592 Plasmid p
C 43	67.8	9.6	6723	9	ADA09843	Ada09843 PiggyBac/
C 44	67.8	9.6	6723	9	ADA09842	Ada09842 Plasmid p
C 45	67.8	9.6	7560	4	AAC85599	Aac85599 Plasmid p

ALIGNMENTS

RESULT 1
ACA55353
ID ACA55353 standard; DNA; 707 BP.
AC ACA55353;
XX
AC ACA55353;
XX
DT 09-JUN-2003 (first entry)
XX
DE Eukaryotic transformation vector piggyBAC ITR cartridge.
XX
KW PiggyBac; transposon; eukaryotic transformation vector; ds;
KW transformed cell; transformed embryo; transgenic; ITR cartridge.
OS Synthetic.
XX
PN US2002173634-A1.
XX
PD 21-NOV-2002.
XX
PF 30-OCT-2001; 2001US-00001189.
XX
PR 31-OCT-2000; 2000US-0244677P.
PR 01-NOV-2000; 2000US-0244984P.
XX
XX (FRAS/) FRASER M J.
PA (LIXX/) LI X.
PA (BEAM/) BEAM T.
XX (HUA/) HUA-VAN A.
XX
PI Fraser MJ, Li X, Beam T, Hua-Van A;
XX WPI; 2003-352597/33.
XX
XX New DNA molecule in the transposon piggyBac, useful for transferring
PT genes into host cells or embryos for transforming the cells of embryos
PT that can be used in making transgenic organisms.
XX
XX Example 3; Fig 3(C1); 15ipp; English.
XX
XX The invention describes a DNA molecule comprising at least 163
CC consecutive nucleotide base pairs of the 3' terminal region beginning at
CC the 3' terminal base pair, and at least 125 consecutive nucleotide base
CC pairs of the 5' terminal region beginning at the 5' terminal base pair, of

CC the piggyBac molecule. The region extends from the restriction site SacI
 CC to the end of the piggyBac molecule. The DNA molecule in the transposon
 CC piggyBac is useful for transferring genes into host cells or embryos for
 CC transforming the cells of embryos. The transformed cells or embryos are
 CC useful for developing or making transgenic organisms. This sequence
 CC represents a minimal sequence cartridge of the eukaryotic transformation
 CC vector piggyBac
 XX
 SQ Sequence 707 BP; 164 A; 174 C; 208 G; 161 T; 0 U; 0 Other;

```

Query Match      100.0%; Score 707; DB 10; Length 707;
Best Local Similarity 100.0%; Pred. No. 2.6e-213;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCCCATCGTCAATTTTACGACAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
DB 1 GGATCCCATCGTCAATTTTACGACAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60

QY 61 ATCATATCGTGGGTCCTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120
DB 61 ATCATATCGTGGGTCCTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120

QY 121 TCGGGGAGGAAGAACCGCGTCTTTTCCGGAGGTTGAAGCGGCATGGAAGAGTTTG 180
DB 121 TCGGGGAGGAAGAACCGCGTCTTTTCCGGAGGTTGAAGCGGCATGGAAGAGTTTG 180

QY 181 CCGAGGATGACTGCTGCTGCAATTTGAGCGAAGAACGACGTTTACCATGATGATTC 240
DB 181 CCGAGGATGACTGCTGCTGCAATTTGAGCGAAGAACGACGTTTACCATGATGATTC 240

QY 241 GGGAGGTTGGCCATGACGCGCTTTTAAACGGTGAACCTGTTCTCAGGCCACCTGGGATA 300
DB 241 GGGAGGTTGGCCATGACGCGCTTTTAAACGGTGAACCTGTTCTCAGGCCACCTGGGATA 300

QY 301 CCAGTTCTCGCGGCTTTTCCGGACACAGTTCCCGATGCTGAGCCGAGCGCATCAGCA 360
DB 301 CCAGTTCTCGCGGCTTTTCCGGACACAGTTCCCGATGCTGAGCCGAGCGCATCAGCA 360

QY 361 ACCGGAACAATACCGGACGACGCGAAGTCCGCGGCTGCGAGATTAAATGACACGG 420
DB 361 ACCGGAACAATACCGGACGACGCGAAGTCCGCGGCTGCGAGATTAAATGACACGG 420

QY 421 GTGCGGCTGGGATATACGTACGCGAGGACGGGTATCCTGGCTGGATGCGCAGAAAT 480
DB 421 GTGCGGCTGGGATATACGTACGCGAGGACGGGTATCCTGGCTGGATGCGCAGAAAT 480

QY 481 GGACATGATATCCCGTCAAGTTACCCGCGGCGCGCTCGTTCAATTCAGCTTTTGAAC 540
DB 481 GGACATGATATCCCGTCAAGTTACCCGCGGCGCGCTCGTTCAATTCAGCTTTTGAAC 540

QY 541 CCGTGGAGGACGGGACAGACTCGCGTGCATAATGCTTTTACAGCGTGTGAGCAGATGA 600
DB 541 CCGTGGAGGACGGGACAGACTCGCGTGCATAATGCTTTTACAGCGTGTGAGCAGATGA 600

QY 601 AGATGCTCGACACGCTGACAGAACGACGATAGATTAAACCCCTAGAAAGATATCATATTG 660
DB 601 AGATGCTCGACACGCTGACAGAACGACGATAGATTAAACCCCTAGAAAGATATCATATTG 660

QY 661 TGACGTAGCTTAAAGATATCATGCTGTAATTTGACGATGGATGCC 707
DB 661 TGACGTAGCTTAAAGATATCATGCTGTAATTTGACGATGGATGCC 707
  
```

RESULT 2
 ACA55360

ID ACA55360 standard; DNA; 4613 BP.

XX AC ACA55360;

XX AC ACA55360;

XX AC ACA55360;

DE Transformation vector piggyBAC related plasmid pCRII-ITR.

KW PiggyBac; transposon; eukaryotic transformation vector; ds;
 KW transformed cell; transformed embryo; transgenic; plasmid; cyclic;
 KW circular.
 XX Synthetic.
 OS
 XX
 PN US2002173634-A1.
 XX
 PD 21-NOV-2002.
 XX
 PF 30-OCT-2001; 2001US-00001189.
 XX
 PR 31-OCT-2000; 2000US-0244677P.
 PR 01-NOV-2000; 2000US-0244984P.
 XX
 PA (FRAS/) FRASER M J.
 PA (LIXX/) Li X.
 PA (BEAM/) BEAM T.
 PA (HUAV/) HUA-VAN A.
 XX
 PI Fraser MJ, Li X, Beam T, Hua-Van A;
 XX
 DR WPI; 2003-352597/33.
 DR P-ESDB; ABU70357, ABU70358.
 XX
 PT New DNA molecule in the transposon piggyBac, useful for transferring
 PT genes into host cells or embryos for transforming the cells of embryos
 PT that can be used in making transgenic organisms.
 XX
 PS Disclosure; Fig 10(B); 151pp; English.
 XX
 CC The invention describes a DNA molecule comprising at least 163
 CC consecutive nucleotide base pairs of the 3' terminal region beginning at
 CC the 3' terminal base pair, and at least 125 consecutive nucleotide base
 CC pairs of the 5' terminal region beginning at the 5'-terminal base pair, of
 CC the piggyBac molecule. The region extends from the restriction site SacI
 CC to the end of the piggyBac molecule. The DNA molecule in the transposon
 CC piggyBac is useful for transferring genes into host cells or embryos for
 CC transforming the cells of embryos. The transformed cells or embryos are
 CC useful for developing or making transgenic organisms. This sequence
 CC represents plasmid used in the creation of minimal sequence eukaryotic
 CC transformation vector piggyBac
 XX
 SQ Sequence 4613 BP; 1064 A; 1183 C; 1273 G; 1093 T; 0 U; 0 Other;

```

Query Match      100.0%; Score 707; DB 10; Length 4613;
Best Local Similarity 100.0%; Pred. No. 6.7e-213;
Matches 707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCCCATCGTCAATTTTACGACAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
DB 294 GGATCCCATCGTCAATTTTACGACAGACTATCTTTCTAGGGTTAATCTAGCTGCATCAGG 353

QY 61 ATCATATCGTGGGTCCTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120
DB 354 ATCATATCGTGGGTCCTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 413

QY 121 TCGGGGAGGAAGAACCGCGTCTTTTCCGGAGGTTGAAGCGGCATGGAAGAGTTTG 180
DB 414 TCGGGGAGGAAGAACCGCGTCTTTTCCGGAGGTTGAAGCGGCATGGAAGAGTTTG 473

QY 181 CCGAGGATGACTGCTGCTGCAATTTGACGTAACCGTAACCGTAACCGTAACCGTAACCGTA 240
DB 474 CCGAGGATGACTGCTGCTGCAATTTGACGTAACCGTAACCGTAACCGTAACCGTA 533

QY 241 GGGAGGTTGGCCATGACGCGCTTTTAAACGGTGAACCTGTTCTCAGGCCACCTGGGATA 300
DB 534 GGGAGGTTGGCCATGACGCGCTTTTAAACGGTGAACCTGTTCTCAGGCCACCTGGGATA 593

QY 301 CCAGTTCTCGCGGCTTTTCCGGACACAGTTCCCGATGCTGAGCCGAGCGCATCAGCA 360
DB 594 CCAGTTCTCGCGGCTTTTCCGGACACAGTTCCCGATGCTGAGCCGAGCGCATCAGCA 653
  
```


OS Synthetic.
 PN US2002173634-A1.
 PD 21-NOV-2002.
 XX 30-OCT-2001; 2001US-00001189.
 XX 31-OCT-2000; 2000US-0244677P.
 PR 01-NOV-2000; 2000US-0244984P.
 XX (FRAS/) FRASER M J.
 PA (LIXX/) LI X.
 PA (BEAM/) BEAM T.
 PA (HUAV/) HUA-VAN A.
 XX Fraser MJ, Li X, Beam T, Hua-Van A;
 XX WPI; 2003-352597/33.
 DR New DNA molecule in the transposon piggyBac, useful for transferring
 PT genes into host cells or embryos for transforming the cells of embryos
 PT that can be used in making transgenic organisms.
 XX Example 11; Fig 13(B); 151pp; English.
 XX The invention describes a DNA molecule comprising at least 163
 CC consecutive nucleotide base pairs of the 3' terminal region beginning at
 CC the 3' terminal base pair, and at least 125 consecutive nucleotide base
 CC pairs of the 5' terminal region beginning at the 5' terminal base pair, of
 CC the piggyBac molecule. The region extends from the restriction site SacI
 CC to the end of the piggyBac molecule. The DNA molecule in the transposon
 CC piggyBac is useful for transferring genes into host cells or embryos for
 CC transforming the cells of embryos. The transformed cells or embryos are
 CC useful for developing or making transgenic organisms. This sequence
 CC represents plasmid used in the creation of minimal sequence eukaryotic
 CC transformation vector piggyBac
 XX Sequence 9009 BP; 2179 A; 2429 C; 2341 G; 2058 T; 0 U; 2 Other;
 SQ
 Query Match 99.8%; Score 705.4; DB 10; Length 9009;
 Best Local Similarity 99.9%; Pred. No. 3e-212;
 Matches 706; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGATCCCATGCTCAATTTTACGACAGACTATCTTCTAGGGTTAATCTAGCTGCATCAGG 60
 DB 8960 GGATCCCATGCTCAATTTTACGACAGACTATCTTCTAGGGTTAATCTAGCTGCATCAGG 8901
 QY 61 ATCATATCGTCGGGTCCTTTTTCGGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 120
 DB 8900 ATCATATCGTCGGGTCCTTTTTCGGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCA 8841
 QY 121 TCGGGGAGAGAACCGCGTCTTTTCCCGGAGTTGAAGCGGCATGGAAGAGTTTG 180
 DB 8840 TCGGGGAGAGAACCGCGTCTTTTCCCGGAGTTGAAGCGGCATGGAAGAGTTTG 8781
 QY 181 CCGAGGATGACTGCTGCTGCTATGAGTGAAGGAAACGACGCTTACCATGATGATTC 240
 DB 8780 CCGAGGATGACTGCTGCTGCTATGAGTGAAGGAAACGACGCTTACCATGATGATTC 8721
 QY 241 GGGAGAGTGTGGCCATGACGCTTTTAAACGGTGAACCTGTTCTGTCAGGCCACCTGGGATA 300
 DB 8720 GGGAGAGTGTGGCCATGACGCTTTTAAACGGTGAACCTGTTCTGTCAGGCCACCTGGGATA 8661
 QY 301 CAGTTCTGTCGGGCTTTTTCGGGACACAGTTCCCGATGGTCAAGCCGATCAGCA 360
 DB 8660 CAGTTCTGTCGGGCTTTTTCGGGACACAGTTCCCGATGGTCAAGCCGATCAGCA 8601
 QY 361 ACCCGAACAAATACCGCGACAGCCGGAAGTCCGTCGGGCTGTCAGATTATGACACGG 420
 DB 8600 ACCCGAACAAATACCGCGACAGCCGGAAGTCCGTCGGGCTGTCAGATTATGACACGG 8541
 QY 421 GTGCGGCGCTGGGATATTACGTGACGAGGACGGGTATCTCTGGCTGGATGCCGAGAAAT 480

Db 8540 GTGCGGCGCTGGGATATTACGTGACGAGGACGGGTATCTCTGGCTGGATGCCGAGAAAT 8481
 QY 481 GGACATGGATACCCCGTGAAGTTACCGCGCGCGCGCTGTTCAATTCAGCTTTTGAAC 540
 Db 8480 GGACATGGATACCCCGTGAAGTTACCGCGCGCGCGCTGTTCAATTCAGCTTTTGAAC 8421
 QY 541 CCGTGGAGGACGGGCGACACTCGCGGTGCAAAATGTGTTTACAGCGTGTGATGGAGCAGATGA 600
 Db 8420 CCGTGGAGGACGGGCGACACTCGCGGTGCAAAATGTGTTTACAGCGTGTGATGGAGCAGATGA 8361
 QY 601 AGATGCTCGACACGCTGCAGAACACGACGTAGATTAAACCTAGAAAGATAATCATATTG 660
 Db 8360 AGATGCTCGACACGCTGCAGAACACGACGTAGATTAAACCTAGAAAGATAATCATATTG 8301
 QY 661 TGACGTACGTTAAAGATAATCATGCGTAAATTTGACGCATGGGATCC 707
 Db 8300 TGACGTACGTTAAAGATAATCATGCGTAAATTTGACGCATGGGATCC 8254
 RESULT 5
 ACA55363/c
 ID ACA55363 standard; DNA; 9012 BP.
 XX AC ACA55363;
 XX AC (first entry)
 DT 06-JUN-2003
 XX Transformation vector piggyBAC related plasmid p(PZ)-Bac-EGFP.
 DE PiggyBac; transposon; eukaryotic transformation vector; ds;
 XX transformed cell; transformed embryo; transgenic; plasmid; cyclic;
 KW circular.
 XX Synthetic.
 OS US2002173634-A1.
 XX 21-NOV-2002.
 PD 30-OCT-2001; 2001US-00001189.
 XX 31-OCT-2000; 2000US-0244677P.
 PR 01-NOV-2000; 2000US-0244984P.
 XX (FRAS/) FRASER M J.
 PA (LIXX/) LI X.
 PA (BEAM/) BEAM T.
 PA (HUAV/) HUA-VAN A.
 XX Fraser MJ, Li X, Beam T, Hua-Van A;
 XX WPI; 2003-352597/33.
 DR New DNA molecule in the transposon piggyBac, useful for transferring
 XX genes into host cells or embryos for transforming the cells of embryos
 XX that can be used in making transgenic organisms.
 XX Example 11; Fig 14(B); 151pp; English.
 CC The invention describes a DNA molecule comprising at least 163
 CC consecutive nucleotide base pairs of the 3' terminal region beginning at
 CC the 3' terminal base pair, and at least 125 consecutive nucleotide base
 CC pairs of the 5' terminal region beginning at the 5' terminal base pair, of
 CC the piggyBac molecule. The region extends from the restriction site SacI
 CC to the end of the piggyBac molecule. The DNA molecule in the transposon
 CC piggyBac is useful for transferring genes into host cells or embryos for
 CC transforming the cells of embryos. The transformed cells or embryos are
 CC useful for developing or making transgenic organisms. This sequence
 CC represents plasmid used in the creation of minimal sequence eukaryotic
 CC transformation vector piggyBac
 XX Sequence 9012 BP; 2180 A; 2419 C; 2353 G; 2060 T; 0 U; 0 Other;
 XX

Db 1199 GGACATGGATACCCGTCGAGTTACCCGCGG-----CTCGTTCATTACGTTTTTGAAC 1252
 Qy 541 CCGTGGAGACGGGACAGACTCGCGTGCAAAATGTGTTTTACAGCGTGATGGAGCAGATGA 600
 Db 1253 CCGTGGAGACGGGACAGACTCGCGTGCAAAATGTGTTTTACAGCGTGATGGAGCAGATGA 1312
 Qy 601 AGATGCTCGACACGCTGCAGAACACGCGAGCTAGATTAAACCTAGAAAAGATAATCATATTG 660
 Db 1313 AGATGCTCGACACGCTGCAGAACACGCGAGCTAGATTAAACCTAGAAAAGATAATCATATTG 1372
 Qy 561 TGACGTAGCTTAAAGATAATCATCGTAAATTTGACGATGGATCC 707
 Db 1373 TGACGTAGCTTAAAGATAATCATCGTAAATTTGACGATGGATCC 1419

RESULT 7

ACA55367
 ID ACA55367 standard; DNA; 4943 BP.

XX ACA55367;

XX 06-JUN-2003 (first entry)

XX Transformation vector piggyBAC related plasmid pBS-ITR-EGFP.

DE PiggyBac; transposon; eukaryotic transformation vector; ds;
 KW transformed cell; transformed embryo; transgenic; plasmid; cyclic;
 KW circular.

XX Synthetic.

XX US2002173634-A1.

XX 21-NOV-2002.

XX 30-OCT-2001; 2001US-00001189.

XX 31-OCT-2000; 2000US-0244677P.

PR 01-NOV-2000; 2000US-0244984P.

XX (FRAS//) FRASER M J.

PA (LIXX//) LI X.

PA (BEAM//) BEAM T.

PA (HUAV//) HUA-VAN A.

XX Fraser MJ, Li X, Beam T, Hua-Van A;

XX WPI; 2003-352597/33.

XX New DNA molecule in the transposon piggyBac, useful for transferring
 PT genes into host cells or embryos for transforming the cells of embryos
 PT that can be used in making transgenic organisms.

XX Example 6; Fig 18(B); 151pp; English.

XX The invention describes a DNA molecule comprising at least 163
 CC consecutive nucleotide base pairs of the 3' terminal region beginning at
 CC the 3' terminal base pair, and at least 125 consecutive nucleotide base
 CC pairs of the 5' terminal region beginning at the 5' terminal base pair, of
 CC the piggyBac molecule. The region extends from the restriction site SacI
 CC to the end of the piggyBac molecule. The DNA molecule in the transposon
 CC piggyBac is useful for transferring genes into host cells or embryos for
 CC transforming the cells of embryos. The transformed cells or embryos are
 CC useful for developing or making transgenic organisms. This sequence
 CC represents plasmid used in the creation of minimal sequence eukaryotic
 CC transformation vector piggyBac

XX Sequence 4943 BP; 1221 A; 1305 C; 1245 G; 1172 T; 0 U; 0 Other;

Query Match 96.9%; Score 685; DB 10; Length 4943;

Best Local Similarity 99.2%; Pred. No. 6.6e-206;

Matches 701; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

Qy 1 GGATCCCATCGTCAATTTTACGACAGATATCTTTCTAGGGTTAATCTAGCTGCATCAGG 60
 Db 719 GGATCCCATCGTCAATTTTACGACAGATATCTTTCTAGGGTTAATCTAGCTGCATCAGG 778
 Qy 61 ATCATATCTGTCGGGTCTTTTTCGGGTCTAGTATCGCCCAAGCTGCGCTATCTGGCA 120
 Db 779 ATCATATCTGTCGGGTCTTTTTCGGGTCTAGTATCGCCCAAGCTGCGCTATCTGGCA 838
 Qy 121 TCGGGAGGAAGAAGCCGTCCTTTTCCGCGAGGTTGAAGCGGCATGAAAGAGATTG 180
 Db 839 TCGGGAGGAAGAAGCCGTCCTTTTCCGCGAGGTTGAAGCGGCATGAAAGAGATTG 898
 Qy 181 CCGAGGATGATCTGCTGCTGCAITGACGTTGAGCGAAACGACGTTTACCATGATTC 240
 Db 899 CCGAGGATGATCTGCTGCTGCAITGACGTTGAGCGAAACGACGTTTACCATGATTC 958
 Qy 241 GGGAGGTGTGGCATGACGCTTTTAAACGCTGAACCTGTTCTGTCAGGCCACCTGGGATA 300
 Db 959 GGGAGGTGTGGCATGACGCTTTTAAACGCTGAACCTGTTCTGTCAGGCCACCTGGGATA 1018
 Qy 301 CCAGTTCGTCGCGCTTTTCCGGACACAGTTCCGATGTTCCAGCCGAAAGCGCATCAGCA 360
 Db 1019 CCAGTTCGTCGCGCTTTTCCGGACACAGTTCCGATGTTCCAGCCGAAAGCGCATCAGCA 1078
 Qy 361 ACCGGAACAATACCGCGACAGCCGGAACCTGCCGTGCCGTGTCAGATTAAATGACAGCG 420
 Db 1079 ACCGGAACAATACCGCGACAGCCGGAACCTGCCGTGCCGTGTCAGATTAAATGACAGCG 1138
 Qy 421 GTGGGCGCTGGGATATTACGTCAGCGAGACGGGTATCTGCTGATCGCGCAGAAAT 480
 Db 1139 GTGGGCGCTGGGATATTACGTCAGCGAGACGGGTATCTGCTGATCGCGCAGAAAT 1198
 Qy 481 GGACATGGATACCCCGTGAGTTACCCGGCGGCGCTCGTTCATTACGTTTTCGAAC 540
 Db 1199 GGACATGGATACCCCGTGAGTTACCCGGCGG-----CTCGTTCATTACGTTTTCGAAC 1252
 Qy 541 CCGTGGAGACGGGACAGACTCGCGGTGCAAAATGTGTTTTACGCGTGTGAGCAGATGA 600
 Db 1253 CCGTGGAGACGGGACAGACTCGCGGTGCAAAATGTGTTTTACGCGTGTGAGCAGATGA 1312
 Qy 601 AGATGCTCGACACGCTGCAGAACACGCGAGCTAGATTAAACCTAGAAAAGATAATCATATTG 660
 Db 1313 AGATGCTCGACACGCTGCAGAACACGCGAGCTAGATTAAACCTAGAAAAGATAATCATATTG 1372
 Qy 661 TGACGTACGTTAAAGATAATCATCGTAAATTTGACGATGGATCC 707
 Db 1373 TGACGTACGTTAAAGATAATCATCGTAAATTTGACGATGGATCC 1419

RESULT 8

ACA55368
 ID ACA55368 standard; DNA; 4944 BP.

XX ACA55368;

XX 06-JUN-2003 (first entry)

XX Transformation vector piggyBAC related plasmid pBS-ITR-EGFP.

DE PiggyBac; transposon; eukaryotic transformation vector; ds;
 KW transformed cell; transformed embryo; transgenic; plasmid; cyclic;
 KW circular.

XX Synthetic.

XX US2002173634-A1.

XX 21-NOV-2002.

XX 30-OCT-2001; 2001US-00001189.

XX 31-OCT-2000; 2000US-0244677P.

PR 01-NOV-2000; 2000US-0244984P.

XX (FRAS/) FRASER M J.
 PA (LIXX/) LI X.
 PA (BEAM/) BEAM T.
 PA (HUA/) HUA-VAN A.
 XX Fraser MJ, Li X, Beam T, Hua-Van A;
 XX WPI; 2003-352597/33.
 XX
 XX New DNA molecule in the transposon piggyBac, useful for transferring
 PT genes into host cells or embryos for transforming the cells of embryos
 PT that can be used in making transgenic organisms.
 XX
 XX Example 6; Fig 19(B); 151pp; English.
 XX
 XX The invention describes a DNA molecule comprising at least 163
 CC consecutive nucleotide base pairs of the 3' terminal region beginning at
 CC the 3' terminal base pair, and at least 125 consecutive nucleotide base
 CC pairs of the 5' terminal region beginning at the 5' terminal base pair, of
 CC the piggyBac molecule. The region extends from the restriction site SacI
 CC to the end of the piggyBac molecule. The DNA molecule in the transposon
 CC piggyBac is useful for transferring genes into host cells or embryos for
 CC transforming the cells of embryos. The transformed cells or embryos are
 CC useful for developing or making transgenic organisms. This sequence
 CC represents plasmid used in the creation of minimal sequence eukaryotic
 CC transformation vector piggyBac
 XX
 XX Sequence 4944 BP; 1221 A; 1302 C; 1247 G; 1174 T; 0 U; 0 Other;
 SQ
 Query Match 96.4%; Score 681.8; DB 10; Length 4944;
 Best Local Similarity 98.9%; Pred. No. 6.8e-205;
 Matches 699; Conservative 0; Mismatches 2; Indels 6; Gaps 1;
 1 GGCATCCCATCGCTCAATTTTACCGAGACTATCTTCTAGGGTTAATCTAGCTCATCAGG 60
 719 GGCATCCCATCGCTCAATTTTACCGAGACTATCTTCTAGGGTTAATCTAGCTCATCAGG 778
 61 ATCATATCGTCGGGTCTTTTTCGGCTCAGTATCGCCCAAGTCGCGCTATCTGGGCA 120
 779 ATCATATCGTCGGGTCTTTTTCGGCTCAGTATCGCCCAAGTCGCGCTATCTGGGCA 838
 121 TCGGGAGGAGGAAGCCGCTGCTTTTCCCGAGGTTGAAGCGGCATGGAAGAGTTTG 180
 839 TCGGGAGGAGGAAGCCGCTGCTTTTCCCGAGGTTGAAGCGGCATGGAAGAGTTTG 898
 181 CCGAGGATGACTGCTGCTGATTCAGCTTGAAGCGGAAACCGATGATGATGATTC 240
 899 CCGAGGATGACTGCTGCTGATTCAGCTTGAAGCGGAAACCGATGATGATGATTC 958
 241 GGGAGGATGCTGCTGCTGATTCAGCTTGAAGCGGAAACCGATGATGATGATTC 300
 959 GGGAGGATGCTGCTGCTGATTCAGCTTGAAGCGGAAACCGATGATGATGATTC 1018
 301 CCAATTCCTCGCGCTTTTTCGGGACACAGTTTCCGGATGTTGAGCCGCAAGCATCAGCA 360
 1019 CCAATTCCTCGCGCTTTTTCGGGACACAGTTTCCGGATGTTGAGCCGCAAGCATCAGCA 1078
 361 ACCGACAAATACCGGACAGCCGGAACCTGCGGTGCGGTGCGATTAATCAGACGG 420
 1079 ACCGACAAATACCGGACAGCCGGAACCTGCGGTGCGGTGCGATTAATCAGACGG 1138
 421 GTGCGGCGCTGGGATATTACGTGACGAGGACGGGTATCTGCGGTGATGCGCGAGAAAT 480
 1139 GTGCGGCGCTGGGATATTACGTGACGAGGACGGGTATCTGCGGTGATGCGCGAGAAAT 1198
 481 GGACATGGATACCCCGTAGTTTACCCCGCGCGCGCTCGTTTCAATTCAGCTTTTGAAC 540
 1199 GGACATGGATACCCCGTAGTTTACCCCGCGCGCGCTCGTTTCAATTCAGCTTTTGAAC 1252
 541 CCGTGGAGGACGGGACAGCTCGCGGTGCAATGTGTTTACAGCGGTGATGAGCAGATGA 600
 1253 CCGTGGAGGACGGGACAGCTCGCGGTGCAATGTGTTTACAGCGGTGATGAGCAGATGA 1312

QY 601 AGATGCTGCACACGCTGCAGAACACCCAGCTAGATTAAACCTAGAAAGATAATCATATTG 660
 DB 1313 AGATGCTGCACACGCTGCAGAACACCCAGCTAGATTAAACCTAGAAAGATAATCATATTG 1372
 QY 661 TGACGTACGTTAAAGATAATCATGCGTAAATTTGACGCATGGATCC 707
 DB 1373 TGACGTACGTTAAAGATAATCATGCGTAAATTTGACGCATGGATCC 1419
 RESULT 9
 ADP70043
 ID ADP70043 standard; DNA; 48502 BP.
 XX
 XX AC ADP70043;
 DT 26-AUG-2004 (first entry)
 XX
 XX Nucleic acid detection-related bacteriophage lambda DNA SeqID6.
 DE
 XX Nucleic acid detection; nucleic acid amplification; microchip;
 KW isothermal; chip substrate; ds.
 XX
 XX Bacteriophage lambda.
 OS
 XX JP2004154008-A.
 PN
 XX 03-JUN-2004.
 PD
 XX 01-NOV-2002; 2002JP-00320330.
 PF
 XX 01-NOV-2002; 2002JP-00320330.
 PR
 XX (BIKE) EIKEN KAGAKU KK.
 PA
 XX WPI; 2004-445081/42.
 DR
 XX
 XX Detecting nucleic acid such as DNA, cDNA and RNA, involves performing
 PT nucleic acid amplification in microchip under isothermal conditions, and
 PT analyzing obtained amplified product on microchip.
 XX
 XX Example 2; SEQ ID NO 6; 49pp; Japanese.
 PS
 XX This invention relates to a novel method of detecting nucleic acid which
 CC comprises performing nucleic acid amplification in a microchip under
 CC isothermal conditions, and analysing the obtained amplified product on
 CC the microchip. The method is useful for detecting nucleic acid and
 CC enables performing a nucleic acid amplification reaction and its analysis
 CC continuously on the microchip without affecting the chip substrate. The
 CC method enables simple, reliable and cost-effective amplification,
 CC analysis and detection of nucleic acid in a microchip. The invention
 CC enables detection of trace amounts of nucleic acid with amplification
 CC efficiency. The present sequence is that of a bacteriophage lambda DNA
 CC sequence which was used in the exemplification of the invention.
 XX
 XX Sequence 48502 BP; 12334 A; 11362 C; 12820 G; 11986 T; 0 U; 0 Other;
 SQ
 Query Match 82.8%; Score 585.4; DB 12; Length 48502;
 Best Local Similarity 99.8%; Pred. No. 8e-174;
 Matches 586; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGCTCAGTCTCGCCCAAGC 104
 3055 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGCTCAGTCTCGCCCAAGC 3114
 105 TCGCGCTATCTCGGCATCGGGGAGGAGAGCCGCTCTTTTTCGGCGAGGTGAAGCG 164
 3115 TCGCGCTATCTCGGCATCGGGGAGGAGAGCCGCTCTTTTTCGGCGAGGTGAAGCG 3174
 165 GCATGGAAGAGGTTTCCGAGGATGACTGCTGCTGATTGACGTTGAGCGGAAACGACG 224
 3175 GCATGGAAGAGGTTTCCGAGGATGACTGCTGCTGATTGACGTTGAGCGGAAACGACG 3234

QY 225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCTTTAAACGGTGAACCTGTTGCTT 284
 Db TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCTTTAAACGGTGAACCTGTTGCTT 3294
 QY 285 CAGGCCACTGGGATACCATGTTGTCGGGGCTTTTCCGGACACAGTTCCGGATGGTGCAGC 344
 Db 3295 CAGGCCACTGGGATACCATGTTGTCGGGGCTTTTCCGGACACAGTTCCGGATGGTGCAGC 3354
 QY 345 CCGAAGCCGATCAGCAACCCGAACAATACCGCGGACAGCCGGAACCTGCGGTGCCGGTGTG 404
 Db 3355 CCGAAGCCGATCAGCAACCCGAACAATACCGCGGACAGCCGGAACCTGCGGTGCCGGTGTG 3414
 QY 405 CAGATTAATGACAGCGGTGCGCGCTGGGATATTACGTCAGCGAGGACGGGTATCTCTGGC 464
 Db 3415 CAGATTAATGACAGCGGTGCGCGCTGGGATATTACGTCAGCGAGGACGGGTATCTCTGGC 3474
 QY 465 TGGATGCCGCGAGAAATGACATGATACCCCGTGTGTTACCGGGCGGGCGCTCGTTTC 524
 Db 3475 TGGATGCCGCGAGAAATGACATGATACCCCGTGTGTTACCGGGCGGGCGCTCGTTTC 3534
 QY 525 ATTACGTTTTTGAACCCGTGGAGCGGACAGCTCGCGGTGCAAAATGTGTTTACAGC 584
 Db 3535 ATTACGTTTTTGAACCCGTGGAGCGGACAGCTCGCGGTGCAAAATGTGTTTACAGC 3594
 QY 585 GTGATGGAGCAGATGAAGATGCTGCACACGCTGCGAGAACACGACGCT 631
 Db 3595 GTGATGGAGCAGATGAAGATGCTGCACACGCTGCGAGAACACGACGCT 3641

RESULT 10

AA544733
 ID AA544733 standard; DNA; 4164 BP.
 AC AA544733;
 XX
 DT 18-DEC-2001 (first entry)
 DE Bacteriophage lambda full-length polynucleotide sequence #158.
 XX
 KW Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
 mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
 cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;
 nervous system disorder; inflammatory disorder; cell differentiation; ds;
 angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;
 genetic disorder; bone regeneration; tendon; ligament; tissue repair;
 cyostatic; antirheumatic; antiarthritic; tendon; ligament; tissue repair;
 antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
 neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic;
 immunostimulant; analgesic; gene therapy.
 XX
 OS Bacteriophage lambda.
 XX
 FN WO200164834-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 26-FEB-2001; 2001WO-US004926.
 XX
 PR 28-FEB-2000; 2000US-00515126.
 PR 18-MAY-2000; 2000US-00577409.
 PR 17-JUN-2000; 2000US-00597707.
 PR 14-JUL-2000; 2000US-00616807.
 PR 19-SEP-2000; 2000US-00664641.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;
 PI Drmanac R;
 XX
 DR WPI; 2001-589862/66.
 DR P-PSDB; AAU27833.
 XX

PT Novel polypeptides and nucleic acids obtained from cDNA libraries
 PT prepared from various human tissues, for diagnosis, treatment of cancer,
 PT neurological, inflammatory disorders and for use in arrays for detection.
 XX Claim 1; SEQ ID NO 158; 153pp; English.
 PS
 XX Sequences AA544576-AA544919 represent full-length polynucleotides and
 CC contig polynucleotides encoding polypeptides of the invention. The DNA
 CC and protein sequences are useful for the treatment, diagnosis and
 CC prevention of various types of disorder in a mammalian subject such as a
 CC human, dog, monkey, mouse, hamster or rat. The disorders include cancers
 CC such as leukaemia, lymphoma and neuroblastoma, autoimmune disorders such
 CC as multiple sclerosis, connective tissue disease, rheumatoid arthritis,
 CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
 CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
 CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
 CC wernicke disease, inflammatory disorders such as nephritis, Crohn's
 CC disease, ischaemia-reperfusion injury, shock, sepsis and inflammation,
 CC bowel disease. The sequences exhibit activity relating to angiogenesis,
 CC cell proliferation, cell differentiation, stem cell growth factor,
 CC activin or inhibin. Therefore, they can be used to manipulate stem cells
 CC in culture to give rise to neuroepithelial cells that can be used to
 CC augment or replace cells damaged by illness, accidental damage or genetic
 CC disorders. The sequences may also be used for regeneration of bone,
 CC cartilage, tendons and ligaments and in tissue repair and burn healing.
 CC Note: Some sequences for this patent did not form part of the printed
 CC specification, but were obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 4164 BP; 906 A; 1123 C; 1307 G; 828 T; 0 U; 0 Other;
 Query Match 82.1%; Score 580.6; DB 4; Length 4164;
 Best Local Similarity 99.3%; Pred. No. 7.8e-173;
 Matches 583; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 104
 Db 178 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 237
 QY 105 TGGCGCTATCTGGCATCGGGAGGAGAGCCGTCCTTTTCCCGCGAGTTGAAGCG 164
 Db 238 TGGCGCTATCTGGCATCGGGAGGAGAGCCGTCCTTTTCCCGCGAGTTGAAGCG 297
 QY 165 GCATGGAAGAGTTTTCGGAGGATGACTGCTGCTGCAATTGACGTCGAGCAAAACGCACG 224
 Db 298 GCATGGAAGAGTTTTCGGAGGATGACTGCTGCTGCAATTGACGTCGAGCAAAACGCACG 357
 QY 225 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCTTTTAAACGGTGAACCTGTTGCTT 284
 Db 358 TTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCTTTTAAACGGTGAACCTGTTGCTT 417
 QY 285 CAGGCCACTGGGATACCATGTTGTCGGGGCTTTTCCGGACACAGTTCCGGATGGTGCAGC 344
 Db 418 CAGGCCACTGGGATACCATGTTGTCGGGGCTTTTCCGGACACAGTTCCGGATGGTGCAGC 477
 QY 345 CCGAAGCCGATCAGCAACCCGAACAATACCGCGGACAGCCGGAACCTGCGGTGCCGGTGTG 404
 Db 478 CCGAAGCCGATCAGCAACCCGAACAATACCGCGGACAGCCGGAACCTGCGGTGCCGGTGTG 537
 QY 405 CAGATTAATGACAGCGGTGCGCGCTGGGATATTACGTCAGCGAGGACGGGTATCTCTGGC 464
 Db 538 CAGATTAATGACAGCGGTGCGCGCTGGGATATTACGTCAGCGAGGACGGGTATCTCTGGC 597
 QY 465 TGGATGCCGCGAGAAATGACATGATACCCCGTGTGTTACCGGGCGGGCGCTCGTTTC 524
 Db 598 TATATGCCGAGAAATGACATGATACCCCGTGTGTTACCGGGCGGGCGCTCGTTTC 657
 QY 525 ATTACGTTTTTGAACCCGTGGAGCGGACAGCTCGCGGTGCAAAATGTGTTTACAGC 584
 Db 658 ATTACGTTTTTGAACCCGTGGAGGACGGGACAGCTCGCGGTGCAAAATGTGTTTACAGC 717
 QY 585 GTGATGGAGCAGATGAAGATGCTGCACACGCTGCGAGAACACGACGCT 631

Db 718 GTGATGGAGCAGATAAAGATGCTCGACACGCTGCGAAGAACGCGAGCT 764

RESULT 11
ABS78924/c
ID ABS78924 standard; DNA; 26565 BP.
XX AC ABS78924;
XX 17-DEC-2002 (first entry)
DT E. coli CFT073 genomic sequence #91.
DE
XX Pathogenic; Escherichia coli CFT073 infection; livestock; pyelonephritis;
KW urinary tract infection; open reading frame; ORF; uropathogenic;
KW antibacterial; aropathic; nephrotropic; gene; ds.
XX Escherichia coli.
OS
XX WO200259320-A2.
PN
XX 01-AUG-2002.
PD
XX 19-OCT-2001; 2001WO-US046833.
PF
XX 19-OCT-2000; 2000US-0242412P.
PR
XX (WISC) WISCONSIN ALUMNI RES FOUND.
PA
XX Blattner FR, Welch RA, Burland VD;
PI
XX WPI; 2002-691532/74.
DR
XX
XX New DNA sequences of the pathogenic Escherichia coli CFT073 strain,
PT useful for preventing or treating E. coli CFT073 infection in humans or
PT livestock.
PT
XX
PS Claim 1; Page 276-291; 765pp; English.
SS
XX The present invention relates to polynucleotide sequences from the genome
CC of the pathogenic Escherichia coli strain CFT073. Almost all the
CC sequences present in E. coli CFT073 are absent in the previously
CC sequenced laboratory strain K-12. The polynucleotide sequences of the
CC invention are useful for preventing, diagnosing or treating E. coli
CC CFT073 infection in humans or livestock. The polynucleotide sequences are
CC useful for preventing urinary tract infections and pyelonephritis.
CC Likewise, the polypeptides encoded by the different open reading frames
CC (ORF1-5) are useful for generating a vaccine against uropathogenic E.
CC coli strains. ABS78834-ABS79085 represent genomic sequences from E. coli
CC strain CFT073
XX
SQ Sequence 26565 BP; 7042 A; 6727 C; 5778 G; 7016 T; 0 U; 2 Other;

Query Match 80.1%; Score 566.2; DB 6; Length 26565;
Best Local Similarity 97.8%; Pred. No. 7.2e-168;
Matches 574; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 45 ATCTAGCTGCATCAGGATCATATGCTCGGGTCTTTTTCGGCTCAGTCATCGCCCAAGC 104
Db ATCCAGCTGCATCAGGATCATATGCTCGGGTCTTTTTCGGCTCAGTCATCGCCCAAGC 3162

Qy 105 TGGCGCTATCTGGGCATCGGGAGGAAGACCGCTGCTTTTCCCGCGAGGTTGAAGCG 164
Db TGGCGCTATCTGGGCATCGGGAGGAAGACCGCTGCTTTTCCCGCGAGGTTGAAGCG 3102

Qy 165 GCATGGAAGAGTTTTCGCGAGGATGACTGCTGCTGCAATTGACGTTGAGCGAAGAACGCAACG 224
Db GCATGGAAGAGTTTTCGCGAGGACGACTGTTGCTGCAATTGACGTTGAGCGAAGAACGCAACG 3042

Qy 225 TTTCACCATGATGATTCGGGAGGTGTCGCATCGACCGCTTTTAACCGTGAACCTGTCGTT 284
Db TTTCACCATGATGATTCGGGAGGTGTCGCATCGACCGCTTTTAACCGTGAACCTGTCGTT 2982

Qy 285 CAGGCCACCTGGGATACCAAGTTTCGTCGCGCTTTTCCGGACACAGTTCCGGATGGTCAGC 344
Db CAGGCCACCTGGGATACCAAGTTTCGTCGCGCTTTTCCGGACACAGTTCCGGATGGTCAGT 2922

Qy 345 CCGAAGCGCATCAGCAACCCGGAACAATACCGGCGACAGCCGGAACCTGCCGTGCCGGTGTG 404
Db CCGAAGCGTATCAGCAACCCGGAACAATACCGGCGACAGCCGGAACCTGCCGTGCCGGTGTG 2862

Qy 405 CAGATTAAATGACAGCGGTGCGCGCTGGGATATTACGTGACGAGGACGGGTATCTCTGGC 464
Db CAGATTAAATGACAGCGGTGCGCGCTGGGATATTACGTGACGAGGACGGGTATCTCTGGC 2802

Qy 465 TGGATCCCGCAGAAATGGACATGATACCCCGTGAGTTACCCGCGCGCGCGCTGCTTC 524
Db TGGATCCCGCAGAAATGGACATGATACCCCGTGAGTTACCCGCGCGCGCGCTGCTTC 2742

Qy 525 ATTACAGTTTTTGAACCCGTTGGAGGACGGGACAGACTCGCGGTGCAAAATGTTTTTACAGC 584
Db ATTACAGTTTTTGAACCCGTTGGAGGACGGGACAGACTCGCGGTGCAAAATGTTTTTACAGC 2682

Qy 585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCGAAGAACGCGAGCT 631
Db GTGATGGAGCAGATGAAGATGCTCGACACGCTGCGAAGAACGCGAGCT 2635

RESULT 12
ADH80491/c
ID ADH80491 standard; DNA; 26565 BP.
XX AC ADH80491;
XX 22-APR-2004 (first entry)
DT Escherichia coli CFT073 genome contig #91.
DE ds; gene; Escherichia coli; CFT073; Escherichia coli CFT073 infection.
XX Escherichia coli; CFT073.
XX US2003165870-A1.
XX 04-SEP-2003.
XX 01-MAR-2002; 2002US-00085959.
XX 01-MAR-2002; 2002US-00085959.
XX (BLAT/) BLATTNER F R.
PA (WELC/) WELCH R A.
PA (BURL/) BURLAND V D.
XX Blattner FR, Welch RA, Burland VD;
XX WPI; 2003-863698/80.
DR New nucleic acid of Escherichia coli CFT073, useful for preparing a
PT composition for diagnosing, treating or preventing infection caused by
PT Escherichia coli CFT073.
XX Claim 1; SEQ ID NO 91; 4pp; English.
XX The invention relates to an isolated Escherichia coli CFT073 nucleic acid
CC molecule. The nucleic acid is useful for preparing a composition for
CC diagnosing, treating or preventing infection caused by Escherichia coli
CC CFT073. The present sequence represents a contig of the Escherichia coli
CC CFT073 genome that is not present in Escherichia coli K-12.
XX
SQ Sequence 26565 BP; 7042 A; 6727 C; 5778 G; 7016 T; 0 U; 2 Other;

Query Match 80.1%; Score 566.2; DB 10; Length 26565;
Best Local Similarity 97.8%; Pred. No. 7.2e-168;
Matches 574; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 45 ATCTAGCTGCATCAGGATCATATCGTCGGGCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
Db |||||
3221 ATCCAGCTGCATCAGGATCATATCGTCGGGCTTTTTTCCGGCTCAGTCATCGCCCAAGC 3162
Qy 105 TGGCGCTATCTGGGCATCGGGAGGAGAAAGCCGCTGCTTTTCCGCGAGGTTGAAGCG 164
Db |||||
3161 TGGCGCTATCTGGGCATCGGGAGGAGAAAGCCGCTGCTTTTCCGCGAGGTTGAAGCG 3102
Qy 165 GCATGGAAGAGTTTCCGAGGATGACCTGCTGCATTGACGTTGACGGAAGACGACG 224
Db |||||
3101 GCATGGAAGAGTTTCCGAGGAGGACTGTTGTCATTTGACGTTGACGGAAGACGACG 3042
Qy 225 TTTACCATGATGATTCGGGAAGGTTGCGCCATGCACGCTTTTAAACGGTGAAGTGTGCTT 284
Db |||||
3041 TTTACCATGATGATTCGGGAAGGTTGCGCCATGCATGCTTTTAAACGGTGAAGTGTGCTT 2982
Qy 285 CAGGCCACCTGGGATACCAAGTTGCTGCGGCTTTTCCGGAACAGATTCGGATGGTTCAGC 344
Db |||||
2981 CAGGCCACCTGGGATACCAAGCCGCTGCGGACTGTTTCCGGAACACAGATTCGGATGGTTCAGT 2922
Qy 345 CCGAAGCGCATCAGCAACCCGAAATACCGCGGACAGCCGGAATGCGGTCGGGTGTG 404
Db |||||
2921 CCGAAGCGCATCAGCAACCCGAAATACCGCGGACAGCCGGAATGCGGTCGGGTGTG 2862
Qy 405 CAGATTAATGACAGCGGTGCGGCTGGGATATTACGTCAGCGAGGACGGGTATCTCTGGC 464
Db |||||
2861 CAGATTAATGACAGCGGTGCGGCTGGGATATTACGTCAGCGAGGACGGGTATCTCTGGC 2802
Qy 465 TGGATGCCGAGAAATGACATGAGTATCCCGTGAGTTTACCGCGGGCGGCTCGTTTC 524
Db |||||
2801 TGGATGCCGAGAAATGACATGAGTATCCCGTGAGTTTACCGCGGGCGGCTCGTTTC 2742
Qy 525 ATTCACGTTTTTGAACCCGTCGAGACCGGCGAGACTCGCGGTGCAAAATGCTTTTACAGC 584
Db |||||
2741 ATTCACGTTTTTGAACCCGTCGAGACCGGCGAGACTCGCGGTGCAAAATGCTTTTACAGC 2682
Qy 585 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCGAGCT 631
Db |||||
2681 GTGATGGAGCAGATGAAGATGCTCGACACGCTGCAGAACACGCGAGCT 2635

RESULT 13
ID ACD19059
XX ACD19059 standard; DNA; 46819 BP.
XX AC ACD19059;
XX AC ACD19059;
DT 27-OCT-2003 (revised)
DT 21-AUG-2003 (first entry)
XX OS Escherichia coli; strain 0157:H7.
XX DE E. coli 0157 unique DNA sequence OZID_72.
XX KW OZID; ds; acute haemorrhagic colitis; haemolytic uraemic syndrome;
KW food poisoning.
XX OS Escherichia coli; strain 0157:H7.
XX PN US2003023075-A1.
XX PD 30-JAN-2003.
XX XX 01-APR-2002; 2002US-00114170.
XX PF 04-DEC-1998; 98US-0110955P.
XX PR 03-DEC-1999; 99US-00453702.
XX XX (BLAT/) BLATTNER F R.
XX PA (BURL/) BURLAND V D.
XX PA (PERN/) PERNA N T.
XX PA (PLUN/) PLUNKETT G.
XX PA (WELC/) WELCH R.
XX XX Blattner FR, Burland VD, Perna NT, Plunkett G, Welch R;

XX WPI; 2003-479497/45.
DR New DNA sequences from Escherichia coli strain 0157:H7, useful for
XX detecting E. coli 0157:H7 in a sample, or in designing diagnostic probes
PT which can be used to distinguish strain 0157:H7 from strain K12 using
PT molecular techniques.
XX Claim 16; SEQ ID NO 72; 33pp; English.
PS The invention relates to an isolated DNA molecule comprising an E. coli
XX strain 0157:H7 sequence selected from a clostridial cytotoxin-like gene,
CC a urasee gene cluster, a RTX toxin-like gene cluster, a locus of
CC enterocyte effacement and 2 genes from its associated lymphocytic phage
CC 933W (a putative serine/threonine kinase and a tail fibre gene). E. coli
CC 0157:H7 can cause food poisoning, specifically acute haemorrhagic colitis
CC (which can develop into haemolytic uraemic syndrome). Also included are
CC an isolated DNA molecule comprising a nucleotide sequence identical to at
CC least 25 contiguous nucleotides contained in DNA sequences selected from
CC ACD18988-ACD19242 (being 255 E.coli 0157 DNA sequences which are not
CC found in E.coli K12), a recombinant DNA construction comprising the DNA
CC above and a method for detecting E. coli 0157:H7 (ATCC 43895) in a sample
CC (or distinguishing between 0157 and K12) using a probe derived from one
CC of the 255 sequences. The DNA sequences are useful in detecting E. coli
CC 0157:H7 in a sample, for the early diagnosis of humans and livestock
CC infected with 0157:H7, and in designing diagnostic probes which can be
CC used to distinguish strain 0157:H7 from strain K12 using molecular
CC techniques. The present sequence is one of the 255 E. coli 0157:H7 DNA
CC sequence (termed OZID_1-OZID255). Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20030023075 (Updated on 27-OCT-2003
CC to standardise OS field)
XX SQ Sequence 46819 BP; 11858 A; 10855 C; 12774 G; 11317 T; 0 U; 15 Other;

Query Match 80.1%; Score 566.2; DB 9; Length 46819;
Best Local Similarity 97.8%; Pred. No. 9.6e-168;
Matches 574; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
Qy 45 ATCTAGCTGCATCAGGATCATATCGTCGGGCTTTTTTCCGGCTCAGTCATCGCCCAAGC 104
Db |||||
14795 ATCCAGCTGCATCAGGATCATATCGTCGGGCTTTTTTCCGGCTCAGTCATCGCCCAAGC 14854
Qy 105 TGGCGCTATCTGGGCATCGGGAGGAGAAAGCCGCTGCTTTTCCCGAGGTTGAAGCG 164
Db |||||
14855 TGGCGCTATCTGGGCATCGGGAGGAGAAAGCCGCTGCTTTTCCCGAGGTTGAAGCG 14914
Qy 165 GCATGGAAGAGTTTCCGAGGATGACCTGCTGCTGCAATTGACGCGAANAACGACG 224
Db |||||
14915 GCATGGAAGAGTTTCCGAGGATGACCTGCTGCTGCAATTGACGCGAANAACGACG 14974
Qy 225 TTTACCATGATGATTCGGGAAGGTTGCGCATGCACGCCCTTTTAAACGGTGAAGTGTGCTT 284
Db |||||
14975 TTTACCATGATGATTCGGGAAGGTTGCGCATGCACGCCCTTTTAAACGGTGAAGTGTGCTT 15034
Qy 285 CAGGCCACCTGGGATACCAAGTTGCTGCGGCTTTTCCGGAACAGATTCGGATGGTTCAGC 344
Db |||||
15035 CAGGCCACCTGGGATACCAAGTCCGCTCCCTCGGACTGTTCCGGAACAGATTCGGATGGTTCAGC 15094
Qy 345 CCGAAGCGCATCAGCAACCCGAAATACCGCGGACAGCCGGAATGCGGTCGGGTGTG 404
Db |||||
15095 CCGAAGCGCATCAGCAACCCGAAATACCGCGGACAGCCGGAATGCGGTCGGGTGTG 15154
Qy 405 CAGATTAATGACAGCGGTGCGGCTGGGATATTACGTCAGCGAGGACGGGTATCTCTGGC 464
Db |||||
15155 CAGATTAATGACAGCGGTGCGGCTGGGATATTACGTCAGCGAGGACGGGTATCTCTGGC 15214
Qy 465 TGGATGCCGAGAAATGACATGAGTATCCCGTGAGTTTACCGCGGGCGGCTCGTTTC 524
Db |||||
15215 TGGATGCCGAGAAATGACATGAGTATCCCGTGAGTTTACCGCGGGCGGCTCGTTTC 15274
Qy 525 ATTCACGTTTTTGAACCCGTCGAGGACCGGCGAGACTCGCGGTGCAAAATGCTTTTACAGC 584

```
Db 15275 ATTCAGCTCTTTAAACCGCTGGAGGACGGGACAGACCGCGGTGCAAAATGTGTTTACAGC 15334
QY 585 GTGATGAGCAGATGAAGATGCTCGACACGCTGCGAGAACACGCGAGCT 631
Db 15335 GTGATGAGCAGATGAAGATGCTCGACACGCTGCGAGAACACGCGAGCT 15381

RESULT 14
ADC00585
ID ADC00585 standard; DNA; 46897 BP.
XX
AC ADC00585;
DT 04-DEC-2003 (first entry)
XX
DE Enterohaemorrhagic E. coli O157:H7-specific nucleic acid SEQ ID NO: 630.
XX ds; gene; enterohaemorrhagic; anti-bacterial.
XX Escherichia coli; O157:H7.
XX
XX JP2002355074-A.
XX
PD 10-DEC-2002.
XX
PF 24-JAN-2002; 2002JP-00015959.
XX
PR 24-JAN-2001; 2001JP-00112010.
XX
PA (UYTS-) UNIV TSUKUBA.
XX
XX WPI; 2003-451640/43.
XX
DR Enterohaemorrhagic Escherichia coli O157:H7-specific nucleic acid molecule
PT and a polypeptide and its use, a polypeptide, a vector and a host cell.
XX
PS Claim 2; SEQ ID NO 630; 2067pp; Japanese.
XX
XX The invention relates to a novel enterohaemorrhagic Escherichia coli
CC O157:H7-specific nucleic acid molecule. A polynucleotide of the invention
CC has anti-bacterial activity. The polypeptide can be used in detection
CC and/or treatment of O157:H7 infection. The nucleotide sequence of the
CC genome of Enterohaemorrhagic E. coli O157:H7 was determined. The present
CC sequence represents an E. coli O157:H7-specific nucleic acid of the
CC invention.
XX
SQ Sequence 46897 BP; 11872 A; 10948 C; 12731 G; 11346 T; 0 U; 0 Other;

Query Match 79.9%; Score 564.6; DB 10; Length 46897;
Best Local Similarity 97.6%; Pred. No. 3.1e-167;
Matches 573; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 45 ATTCAGTCATCAGATCATATATCGTCGGGCTCTTTTCGGCTCAGTCATCGCCCAAGC 104
Db 14795 ATCCAGTCGATCAGGATCATATATCGTCGGGCTCTTTTCGGCTCAGTCATCGCCCAAGC 14854
QY 105 TGGCGCTATCTGGCATCGGGAGGAGAGCCGTCCTTTCCCGGAGTTGAAGC 164
Db 14855 TGGCGCTATCTGGCATCGGGAGGAGAGCCGTCCTTTTCGGCTCAGTCATCGCCCAAGC 14914
QY 165 GCATGAAAGAGTTTCCGAGGATGACTGCTGCTGCAATTGACCTTGAGCGAAGAACGACG 224
Db 14915 GCATGAAAGAGTTTCCGAGGATGACTGCTGCTGCAATTGACCTTGAGCGAAGAACGACG 14974
QY 225 TTATCAGATGATTCGGGAAGGTGGCCATGACGCCCTTTAACCGTGAACCTGTCGTT 284
Db 14975 TTATCAGATGATTCGGGAAGGTGGCCATGACGCCCTTTAACCGTGAACCTGTCGTT 15034
QY 285 CAGGCCACCTGGATACCACTGTCGCGGCTTTTCGGACAGATTCGGGATGTCAGC 344
Db 15035 CAGGCCACCTGGATACCACTGTCGCGGCTTTTCGGACAGATTCGGGATGTCAGC 15094
```

RESULT 15

ACD19238/C

ID ACD19238 standard; DNA; 22306 BP.

XX

AC ACD19238;

XX

DT 27-OCT-2003 (revised)

DT 21-AUG-2003 (first entry)

XX

DE E. coli O157 unique DNA sequence OZID_251.

XX

KW OZID; ds; acute haemorrhagic colitis; haemolytic uraemic syndrome;
food poisoning.

XX

OS Escherichia coli; strain O157:H7.

XX

PN US2003023075-A1.

XX

PD 30-JAN-2003.

XX

PF 01-APR-2002; 2002US-00114170.

XX

PR 04-DEC-1998; 98US-0110955P.

PR

03-DEC-1999; 99US-00453702.

XX

(BLAT/) BLATTNER F R.

PA (BURL/) BURLAND V D.

PA (PERN/) PERNA N T.

PA (PLUN/) PLUNKETT G.

PA (WELC/) WELCH R.

XX

PI Blattner FR, Burland VD, Perna NT, Plunkett G, Welch R;

XX

WPI; 2003-479497/45.

XX

New DNA sequences from Escherichia coli strain O157:H7, useful for
detecting E. coli O157:H7 in a sample, or in designing diagnostic probes
which can be used to distinguish strain O157:H7 from strain K12 using
molecular techniques.

XX

PS Claim 16; SEQ ID NO 251; 33pp; English.

XX

CC The invention relates to an isolated DNA molecule comprising an E. coli
strain O157:H7 sequence selected from a clostridial cytotoxin-like gene,
a urease gene cluster, a RTX toxin-like gene cluster, a locus of
enterocyte effacement and 2 genes from its associated lymphocytic phage
933W (a putative serine/threonine kinase and a tail fibre gene). E. coli
O157:H7 can cause food poisoning, specifically acute haemorrhagic colitis
(which can develop into haemolytic uraemic syndrome). Also included are
an isolated DNA molecule comprising a nucleotide sequence identical to at
least 25 contiguous nucleotides contained in DNA sequences selected from
ACD18988-ACD19242 (being 255 E.coli O157 DNA sequences which are not

CC found in E.coli K12), a recombinant DNA construction comprising the DNA
CC above and a method for detecting E. coli O157:H7 (ATCC 43895) in a sample
CC (or distinguishing between O157 and K12) using a probe derived from one
CC of the 255 sequences. The DNA sequences are useful in detecting E. coli
CC O157:H7 in a sample, for the early diagnosis of humans and livestock
CC infected with O157:H7, and in designing diagnostic probes which can be
CC used to distinguish strain O157:H7 from strain K12 using molecular
CC techniques. The present sequence is one of the 255 E. coli O157:H7 DNA
CC sequence (termed O2ID 1-O2ID255). Note: The sequence data for this patent
CC did not form part of the printed specification, but was obtained in
CC electronic format directly from the USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20030023075 (Updated on 27-OCT-2003
CC to standardise OS field)
XX
XX

SQ	Sequence	22306 BP; 4744 A; 6842 C; 5512 G; 5192 T; 0 U; 16 Other;
Query Match	79.4%;	Score 561.4; DB 9; Length 22306;
Best Local Similarity	97.3%;	Pred. No. 2.2e-166;
Matches	571; Conservative	0; Mismatches 16; Indels 0; Gaps 0;
QY	45	ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGCTCAGTCATCGCCCAAGC 104
DB	17193	ATCCAGTTGCATCAGGATCATATCGTCGGGTCTTTTTCGGCTCAGTCATCGCCCAAGC 17134
QY	105	TGGCGCTATCTGGGCATCGGGAGGAAGACCCGTCGCTTTTCCCGCAGGTTGAAGCG 164
DB	17133	TGGCGCTATCTGGGCATCGGGAGGAAGACCCGTCGCTTTTCCCGCAGGTTGAAGCG 17074
QY	165	GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGCATTTGACGCGAAACGCGACG 224
DB	17073	GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGCATTTGACGCGAAACGCGACG 17014
QY	225	TTTACCATGATGATTTCGGGAAGGTGTGCCATGCACGCTTTTAAACGGTGAACCTGTTGTT 284
DB	17013	TTTACCATGATGATTTCGGGAAGGTGTGCCATGCACGCTTTTAAACGGTGAACCTGTTGTT 16954
QY	285	CAGGCCACCTGGGATACCAAGTTGCTGCGGGCTTTTTCGGACACAGTTCCGGATGGTCAGC 344
DB	16953	CAGGCCACCTGGGATACCCGTCCTCCGACTGTTTCGGACACAGTTCCGGATGGTCAGC 16894
QY	345	CCGAAGCGCATCAGCAACCCGAACAAATACCGCGGACAGCCGGAACTGCCGTGCCGGTGTG 404
DB	16893	CCGAAGCGCATCAGCAACCCGAACAAATACCGCGGACAGCCGGAACTGCCGTGCCGGTGTG 16834
QY	405	CAGATTAAATGACAGCGGTCCGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 464
DB	16833	CAGATTAAATGACAGCGGTCCGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGC 16774
QY	465	TGGATGCCCGCAGAAATGGACATGGATACCCCGTGAGTTTACCCGGCGGCGCCCTCGTTTC 524
DB	16773	TGGATGCCCGCAGAAATGGACATGGATACCCCGTGAGTTTACCCGGCGGCGCCCTCGTTTC 16714
QY	525	ATTACAGTTTGTGAACCCGTGGAGACCGGGCAGACTCCGGGTGCAAAATGCTGTTTACAGC 584
DB	16713	ATTACAGTCTTTTGAACCCGTGGAGACCGGGCAGACTCCGGGTGCAAAATGCTGTTTACAGC 16654
QY	585	GTGATGGAGCAGATGAAGATGCTGCACACGCTGCAGAAACACGACGCT 631
DB	16653	GTGATGGAGCAGATGAAGATGCTGCACACGCTGCAGAAACACGACGCT 16607

Search completed: October 12, 2005, 15:19:58
Job time : 3479 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 12, 2005, 14:11:50 ; Search time 27480 Seconds
(without alignments)
979.310 Million cell updates/sec

Title: US-10-826-523-40
Perfect score: 707
Sequence: 1 ggaatccatgcgtcaatttt.....aaaattgacgcatggatccc 707

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gss1: *
9: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	585.4	82.8	682	7	CK781302
C 2	585.4	82.8	703	7	CF537771
C 3	585.4	82.8	711	7	CF743678
C 4	585.4	82.8	853	9	CF65566
C 5	584.4	82.7	703	5	BQ154655
C 6	584.4	82.7	719	6	CD350897
C 7	583.8	82.6	751	9	CR087413
C 8	580.4	82.1	716	9	CR131675
C 9	580.4	82.0	786	9	CR108810
C 10	565.4	80.0	691	6	CD350776
C 11	564	79.8	762	9	CR035207
C 12	563.4	79.7	749	6	CD351273
C 13	562	79.5	733	9	CR139473
C 14	555	78.5	633	9	CR167678
C 15	554.4	78.4	791	6	CB520716
C 16	550.4	77.9	630	9	CR013093
C 17	549	77.7	702	9	CR084987
C 18	545.4	77.1	611	4	B1423083
C 19	543.4	76.9	716	7	CK781284
C 20	540.8	76.5	716	7	CF851544
C 21	537.2	76.0	706	1	AV731514
C 22	535.4	75.7	730	9	CR077673
C 23	525	74.3	635	9	EX982794
C 24	521.8	73.8	536	9	CR026633

ALIGNMENTS

RESULT 1
LOCUS CK781302/c
DEFINITION CK781302 682 bp mRNA linear EST 23-FEB-2004
IMAGE:30619060 5', mRNA sequence.

ACCESSION CK781302
VERSION CK781302
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 682)
TITLE NIH-MGC http://mgc.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-f@mail.nih.gov

Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pYX-5.
Location/Qualifiers
1. .682

FEATURES
source
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30619060"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (tr phage resistant)"
/clone_lib="NIH BMAP GI0"
/notes="Organ: Brain; Vector: pYX- Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag

REFERENCE 1 (bases 1 to 719)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 JOURNAL
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: rsb@nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.
 Location/Qualifiers
 1. 719
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6853243"
 /tissue_type="whole brain"
 /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
 /lab_host="DH10B (TI phage resistant)"
 /clone_lib="NIH BMAP G10"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according
 Bonaïdo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is AGCGAGCAG. This library was created for the University
 Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

ORIGIN

Query Match 82.7%; Score 584.4; DB 6; Length 719;
 Best Local Similarity 99.7%; Pred. No. 4.8e-164;
 Matches 585; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 104
 DB 663 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 604
 QY 105 TGGCGCTATCTGGGCATCGGGAGGAAGAGCCGCTCTTTTCCCGGAGGTTGAAGCG 164
 DB 603 TGGCGCTATCTGGGCATCGGAGGAAGAGCCGCTCTTTTCCCGGAGGTTGAAGCG 544
 QY 165 GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGCTATGACGTTGAGCGAAACGCAAG 224
 DB 543 GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGCTATGACGTTGAGCGAAACGCAAG 484
 QY 225 TTTACCATGATGATTCGGGAAGGTGGCCATGACGCCCTTTTAAACGGTGAACCTGTCGTT 284
 DB 483 TTTACCATGATGATTCGGGAAGGTGGCCATGACGCCCTTTTAAACGGTGAACCTGTCGTT 424
 QY 285 CAGGCCACCTGGGATACAGTTCTGTCGGCTTTTCCGGACACAGTTTCCGGATGGTCAGC 344
 DB 423 CAGGCCACCTGGGATACAGTTCTGTCGGCTTTTCCGGACACAGTTTCCGGATGGTCAGC 364
 QY 345 CCGAAGCGCATCAGCAACCCGAAACAATACCGGGGACAGCGGAACTGCGCGGTGTG 404
 DB 363 CCGAAGCGCATCAGCAACCCGAAACAATACCGGGGACAGCGGAACTGCGCGGTGTG 304
 QY 405 CAGATTAAATGACAGCGGTGGGCTGGGATATTACGTACGGAGGACGGGTATCTCTGGC 464

Db 303 CAGATTAAATGACAGCGGTGGGCTGGGATATTACGTACGGAGGACGGGTATCTCTGGC 244
 QY 465 TGGATGCGCAGAAATGGACATGATACCCCGTAGTTACCCGCGGCGCGCTCGTTC 524
 Db 243 TGGATGCGCAGAAATGGACATGATACCCCGTAGTTACCCGCGGCGCGCTCGTTC 184
 QY 525 ATTACAGTTTGAACCCGTGGAGGAGGCGCAGACTCGCGGTGCAAAATGTTTACAGC 584
 Db 183 ATTACAGTTTGAACCCGTGGAGGAGGCGCAGACTCGCGGTGCAAAATGTTTACAGC 124
 QY 585 GTCATGAGCAGATGAAGATGCTCGACAGCTCAGAAACACGAGCT 631
 Db 123 GTCATGAGCAGATGAAGATGCTCGACAGCTCAGAAACACGAGCT 77

RESULT 7

CR087413/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

Location/Qualifiers

1. 751

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/clone="MHPP438f24"

/clone_lib="MHPP"

ORIGIN

Query Match 82.6%; Score 583.8; DB 9; Length 751;
 Best Local Similarity 99.7%; Pred. No. 7.3e-164;
 Matches 585; Conservative 0; Mismatches 2; Indels 0; Gaps 0;QY 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 104
 DB 623 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 564
 QY 105 TGGCGCTATCTGGGCATCGGGAGGAAGAGCCGCTCTTTTCCCGGAGGTTGAAGCG 164
 DB 563 TGGCGCTATCTGGGCATCGGGAGGAAGAGCCGCTCTTTTCCCGGAGGTTGAAGCG 504
 QY 165 GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGCTATGACGTTGAGCGAAACGCAAG 224
 DB 503 GCATGGAAGAGTTTCCGAGGATGACTGCTGCTGCTATGACGTTGAGCGAAACGCAAG 444
 QY 225 TTTACCATGATGATTCGGGAAGGTGGCCATGACGCCCTTTTAAACGGTGAACCTGTCGTT 284
 DB 443 TTTACCATGATGATTCGGGAAGGTGGCCATGACGCCCTTTTAAACGGTGAACCTGTCGTT 384
 QY 285 CAGGCCACCTGGGATACAGTTCTGTCGGCTTTTCCGGACACAGTTTCCGGATGGTCAGC 344
 DB 383 CAGGCCACCTGGGATACAGTTCTGTCGGCTTTTCCGGACACAGTTTCCGGATGGTCAGC 324
 QY 345 CCGAAGCGCATCAGCAACCCGAAACAATACCGGGGACAGCGGAACTGCGCGGTGTG 404
 DB 323 CCGAAGCGCATCAGCAACCCGAAACAATACCGGGGACAGCGGAACTGCGCGGTGTG 264
 QY 405 CAGATTAAATGACAGCGGTGGGCTGGGATATTACGTACGGAGGACGGGTATCTCTGGC 464

```
Db      263  CAGATTAAATGACAGCGGTGCGGCGCTGGGATATATACGTACGCGAGGACGGGTATCCTGGC 204
Qy      465  TGGATGCCGAGAAATGACATGATACCCCGTGAATACCCGCGGCGCGCTCGTTC 524
Db      203  TGGATGCCGAGAAATGACATGATACCCCGTGAATACCCGCGGCGCGCTCGTTC 144
Qy      525  ATTACGCTTTTGAACCCGTTGAGAGGACGGGAGACTCGCGGTGCAAAATGTGTTTACAGC 584
Db      143  ATTACGCTTTTGAACCCGTTGAGAGGACGGGAGACTCGCGGTGCAAAATGTGTTTACAGC 84
Qy      585  GTGATGGAGCAGATGAAGATGCTGCACACGCTGCAGAACGCGCAGCT 631
Db      83  GTGATGGAGCAGATGAAGATGCTGCACACGCGAGCAGAACGCGAGCT 37

RESULT 8
CR131675/c
LOCUS   CR131675               716 bp      DNA      linear      GSS 06-JUL-2004.
DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and
            chromosome engineering clone MHP438f23, genomic survey sequence.
ACCESSION CR131675
VERSION   CR131675.1 GI:49879128
KEYWORDS  GSS; genome survey sequence; MICER.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
REFERENCE 1 (bases 1 to 716)
AUTHORS  Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
Rogers,J. and Bradley,A.
TITLE     Direct Submission
JOURNAL   Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES  Location/Qualifiers
            source
            1..716
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            /clone="MHP438f23"
            /clone_lib="MHPP"

ORIGIN
Query Match      82.1%; Score 580.4; DB 9; Length 716;
Best Local Similarity 99.8%; Pred. No. 7,6e-163;
Matches 581; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      45  ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGCTCAGTCATCGCCCAAGC 104
Db      582  ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGCTCAGTCATCGCCCAAGC 523
Qy      105  TGGCGCTATCTGGGCATCGGGAGGAGAGAGCCGCTGCTTTTCCGCGAGGTTGAAGCG 164
Db      532  TGGCGCTATCTGGGCATCGGGAGGAGAGAGCCGCTGCTTTTCCGCGAGGTTGAAGCG 463
Qy      165  GCATGGAAAGAGTTTCCGAGGATGACTGCTGCTGCATTGAGCGTTGAGCGAAACGCACG 224
Db      462  GCATGGAAAGAGTTTGCAGGATGACTGCTGCTGCATTGAGCGTTGAGCGAAACGCACG 403
Qy      225  TTTACCATGATGATTCGGGAAGGTGTGGCATGCAAGCGCTTTTAAACGGTGAACCTGCTT 284
Db      402  TTTACCATGATGATTCGGGAAGGTGTGGCATGCAAGCGCTTTTAAACGGTGAACCTGCTT 343
Qy      285  CAGGCCACTGGGATACAGTTCGTCGGGCTTTTCCGGACACAGTTCGGATGTGACG 344
Db      342  CAGGCCACTGGGATACAGTTCGTCGGGCTTTTCCGGACACAGTTCGGATGTGACG 283
Qy      345  CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCCGAACTGCGCGTGTG 404
Db      282  CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACAGCCGAACTGCGCGTGTG 223
Qy      405  CAGATTAAATGACAGCGGTGCGGCGCTGGGATATATACGTACGCGAGGACGGGTATCCTGGC 464
```

```
Db      222  CAGATTAAATGACAGCGGTGCGGCGCTGGGATATATACGTACGCGAGGACGGGTATCCTGGC 163
Qy      465  TGGATGCCGAGAAATGACATGATACCCCGTGAATACCCGCGGCGCGCTCGTTC 524
Db      162  TGGATGCCGAGAAATGACATGATACCCCGTGAATACCCGCGGCGCGCTCGTTC 103
Qy      525  ATTACGCTTTTGAACCCGTTGAGAGGACGGGAGACTCGCGGTGCAAAATGTGTTTACAGC 584
Db      102  ATTACGCTTTTGAACCCGTTGAGAGGACGGGAGACTCGCGGTGCAAAATGTGTTTACAGC 43
Qy      585  GTGATGGAGCAGATGAAGATGCTGCACACGCTGCAGAACGCGCAGC 626
Db      42  GTGATGGAGCAGATGAAGATGCTGCACACGCTGCAGAACAGC 1

RESULT 9
CR108810/c
LOCUS   CR108810               786 bp      DNA      linear      GSS 05-JUL-2004
DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and
            chromosome engineering clone MHP22016, genomic survey sequence.
ACCESSION CR108810
VERSION   CR108810.1 GI:49856225
KEYWORDS  GSS; genome survey sequence; MICER.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
REFERENCE 1 (bases 1 to 786)
AUTHORS  Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
Rogers,J. and Bradley,A.
TITLE     Direct Submission
JOURNAL   Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES  Location/Qualifiers
            source
            1..786
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            /clone="MHP22016"
            /clone_lib="MHPP"

ORIGIN
Query Match      82.0%; Score 580; DB 9; Length 786;
Best Local Similarity 100.0%; Pred. No. 1e-162;
Matches 580; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      52  TGCATCAGGATCATATCGTCGGGTCTTTTTCGGCTCAGTCATCGCCCAAGCTGGCGCT 111
Db      786  TGCATCAGGATCATATCGTCGGGTCTTTTTCGGCTCAGTCATCGCCCAAGCTGGCGCT 727
Qy      112  ATCTGGGCATCGGGAGGAGAGAGCCGCTGCTTTTCCGCGAGGTTGAAGCGCATGGA 171
Db      726  ATCTGGGCATCGGGAGGAGAGAGCCGCTGCTTTTCCGCGAGGTTGAAGCGCATGGA 667
Qy      172  AAGAGTTTGGCGAGGATGACTGCTGCTGCATTGAGCGTTGAGCGAAACGCACGTTTACCA 231
Db      666  AAGAGTTTGGCGAGGATGACTGCTGCTGCATTGAGCGTTGAGCGAAACGCACGTTTACCA 607
Qy      232  TGATGATTCGGGAAGGTGTGGCCATGCACGCTTTTAAACGGTGAACCTGTTCCGTCAGGCCA 291
Db      606  TGATGATTCGGGAAGGTGTGGCCATGCACGCTTTTAAACGGTGAACCTGTTCCGTCAGGCCA 547
Qy      292  CCTGGGATACAGTTCGTCGCGGCTTTTCCGGACACAGTTCGGATGTGACGCCGAAGC 351
Db      546  CCTGGGATACAGTTCGTCGCGGCTTTTCCGGACACAGTTCGGATGTGACGCCGAAGC 487
Qy      352  GCATCAGCAACCCGAAACAATACCGGCGACAGCCGAACTGCGCGTGTGCGAGATTA 411
Db      486  GCATCAGCAACCCGAAACAATACCGGCGACAGCCGAACTGCGCGTGTGCGAGATTA 427
Qy      412  ATGACAGCGGTGCGGCGCTGGGATATATACGTACGCGAGGACGGGTATCCTGGCTGGATGC 471
```

Db 426 ATGACACGGTGGCGGCTCGGATATTAACGTACGACGAGGACGGGTATCCCTGGCTGGATGC 367
 Qy 472 CGCAGAAATGACATGATACCCGTGAGTATACCCGGCGGCGCGCTCGTTCATTACAG 531
 Db 366 CGCAGAAATGACATGATACCCGTGAGTATACCCGGCGGCGCGCTCGTTCATTACAG 307
 Qy 532 TTTTGAACCGTGGAGGACGGGACACTCGCGTGCAGATGTTTACAGCGTGTATGG 591
 Db 306 TTTTGAACCGTGGAGGACGGGACACTCGCGTGCAGATGTTTACAGCGTGTATGG 247
 Qy 592 AGCAGATGAAGATGCTCGACACGCTGCAGAACACGACGCT 631
 Db 246 AGCAGATGAAGATGCTCGACACGCTGCAGAACACGACGCT 207

RESULT 10
 CD350776/c
 LOCUS
 DEFINITION
 UI-M-G10-cgh-g-19-0-UI.r1 NIH_BMAP_G10 Mus musculus CDNA clone
 IMAGE:6853388 5', mRNA sequence.
 CD350776
 CD350776.1 GI:31142363
 EST.
 Mus musculus (house mouse)
 ORGANSIM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 691)
 NTH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-x@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
 http://genome.uiowa.edu/distribution/mousefl.html
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pyx-5.

Location/Qualifiers

FEATURES

source
 1. 691
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6853388"
 /tissue_type="whole brain"
 /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_G10"
 /notes="Organ: Brain; Vector: pyx- Asc; Site 1: EcoR I;
 Site 2: Not I; The library was constructed according
 to the protocol of Benton and Davis (1978). The library was
 constructed from total RNA isolated from E13.5, E14.5, E16.5,
 E17.5 embryos, and was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with Not I and then cloned
 directionally into pyx-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is AGCAGACAG. This library was created for the University
 of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

ORIGIN

Query Match 80.0%; Score 565.4; DB 6; Length 691;

Best Local Similarity 99.8%; Pred. No. 2.4e-158;

Matches 566; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 45 ATCTAGCTGCATCAGATCATATCGTGGGTCTTTTTCGGCTCAGTCAATCGCCCAAGC 104
 Db 567 ATCCAGCTGCATCAGGATCATATCGTGGGTCTTTTTCGGCTCAGTCAATCGCCCAAGC 508
 Qy 105 TGGCGCTATCTGGGATCGGGGAGGAGCCGTCCTTTTCCCGCGAGGTTGAAGCG 164
 Db 507 TGGCGCTATCTGGGATCGGGGAGGAGCCGTCCTTTTCCCGCGAGGTTGAAGCG 448
 Qy 165 GCATGAAAGAGTGTTCGCCGAGGATGACTGCTGCTGATTGACCTTTTAAACGGTGAACCTGTCGTT 224
 Db 447 GCATGAAAGAGTGTTCGCCGAGGATGACTGCTGCTGATTGACCTTTTAAACGGTGAACCTGTCGTT 388
 Qy 225 TTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCCTTTTAAACGGTGAACCTGTCGTT 284
 Db 387 TTTTACCATGATGATTCGGGAAGGTGTGGCCATGCACGCCCTTTTAAACGGTGAACCTGTCGTT 328
 Qy 285 CAGGCCACCTGGGATACCACTGCTCGCGCTTTTCCGGACACAGTTCGGGATGGTCAGC 344
 Db 327 CAGGCCACCTGGGATACCACTGCTCGCGCTTTTCCGGACACAGTTCGGGATGGTCAGC 268
 Qy 345 CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACACGCGGAACTGCCCTGCCGCTGTG 404
 Db 267 CCGAAGCGCATCAGCAACCCGAAACAATACCGGCGACACGCGGAACTGCCCTGCCGCTGTG 208
 Qy 405 CAGATTAAATGACAGCGGTGGCGCTGGGATATTAAGTTCAGCAGGACGGGTATCTCGGC 464
 Db 207 CAGATTAAATGACAGCGGTGGCGCTGGGATATTAAGTTCAGCAGGACGGGTATCTCGGC 148
 Qy 465 TGGATCCCGCAGAAATGGACATGATACCCGCTGAGTATACCCGCGGCGCGCTCGTTC 524
 Db 147 TGGATCCCGCAGAAATGGACATGATACCCGCTGAGTATACCCGCGGCGCGCTCGTTC 88
 Qy 525 ATTACAGCTTTTGAACCCGCTGGAGGACGGGACAGACTCGCGGTGCAAAATGTGTTTACAGC 584
 Db 87 ATTACAGCTTTTGAACCCGCTGGAGGACGGGACAGACTCGCGGTGCAAAATGTGTTTACAGC 28
 Qy 585 GTGATCGAGCAGATGAAGATGCTCGAC 611
 Db 27 GTGATCGAGCAGATGAAGATGCTCGAC 1

RESULT 11

CR035207

LOCUS

DEFINITION

Reverse strand read from insert in 5'HPRT insertion targeting and

chromosome engineering clone MHPN278n13, genomic survey sequence.

CR035207

VERSION

GSS; genome survey sequence; MICER.

KEYWORDS

Mus musculus (house mouse)

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 762)

ADAMS, D.J., BIGGS, P.J., COX, A.V., DAVIES, R.M., VAN DER WEYDEN, L.,

JONKERS, J., SMITH, J., PLUMB, R.W., TAYLOR, R.G., NISHIJIMA, I., YU, Y.,

ROGERS, J. and BRADLEY, A.

Direct Submission

Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. http://www.sanger.ac.uk/MICER

FEATURES

Location/Qualifiers

source

1. 762

/organism="Mus musculus"

/mol_type="genomic DNA"

/db_xref="taxon:10090"

/clone="MHPN278n13"

/clone_lib="MHPN"

ORIGIN

Query Match 79.8%; Score 564; DB 9; Length 762;

Best Local Similarity 100.0%; Pred. No. 6.5e-158;

Matches 564; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 68 CGTGGGCTCTTTTCCGGCTCAGTCATCGCCCAAGCTGGGCTATCTGGGCATCGGGGA 127
Db 2 CGTGGGCTCTTTTCCGGCTCAGTCATCGCCCAAGCTGGGCTATCTGGGCATCGGGGA 61
Qy 128 GGAAGAAGCCGCTGCTTTTCCCGAGAGTTGAAGCGGCATGGAAGAGTTTGGCAGGA 187
Db 62 GGAAGAAGCCGCTGCTTTTCCCGAGAGTTGAAGCGGCATGGAAGAGTTTGGCAGGA 121
Qy 188 TGACTGCTGCTGCAATGAGCTGAGCGAAAGCGACGTTTACCATGATGATTCGGGAGG 247
Db 122 TGACTGCTGCTGCAATGAGCTGAGCGAAAGCGACGTTTACCATGATGATTCGGGAGG 181
Qy 248 TGTGGCCATGACAGCGCTTTTAAACGGTGAATCTGTTCCAGGCGACCTGGGATACCAATTC 307
Db 182 TGTGGCCATGACAGCGCTTTTAAACGGTGAATCTGTTCCAGGCGACCTGGGATACCAATTC 241
Qy 308 CTCGCGGCTTTTCCGGACACAGTTCCGGATGTCAGCCCGAAGCGCATCAGCAACCCGAA 367
Db 242 CTCGCGGCTTTTCCGGACACAGTTCCGGATGTCAGCCCGAAGCGCATCAGCAACCCGAA 301
Qy 368 CAATACCGGCGACAGCGGAACTGCGGTGCCGTGTGCAGATTAAATGACAGCGGTGGGC 427
Db 302 CAATACCGGCGACAGCGGAACTGCGGTGCCGTGTGCAGATTAAATGACAGCGGTGGGC 361
Qy 428 GCTGGGATATTACGTCAGCGAGGACGGTATCTTCCGCTGGATGCCGACAGAAATGGACATG 487
Db 362 GCTGGGATATTACGTCAGCGAGGACGGTATCTTCCGCTGGATGCCGACAGAAATGGACATG 421
Qy 488 GATACCCCGTGAATTAACCGCGGCGCGCTGCTGTTCAATTCAGTTTGAACCCGTGGA 547
Db 422 GATACCCCGTGAATTAACCGCGGCGCGCTGCTGTTCAATTCAGTTTGAACCCGTGGA 481
Qy 548 GGACGGGCGAGTCCGCGTGCAATGCTTTTACAGCTGATGGAGCAGATGAGATGCT 607
Db 482 GGACGGGCGAGTCCGCGTGCAATGCTTTTACAGCTGATGGAGCAGATGAGATGCT 541
Qy 608 CGACACGCTGCAGAACACGCGAGCT 631
Db 542 CGACACGCTGCAGAACACGCGAGCT 565
```

```
RESULT 12
CD351273/c
LOCUS UI-M-G10-cgh-f-22-0-UI.r1 NIH_BMAP_G10 Mus musculus cDNA clone
DEFINITION IMAGE:6853367 5', mRNA sequence.
```

```
ACCESSION CD351273.1 GI:31142608
VERSION EST.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
```

```
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mouseefi.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
```

```
Seq primer: pYX-5,
Location/Qualifiers
1..749
/organism="Mus musculus"
/mol_type="mRNA"
```

```
FEATURES
source
```

```
RESULT 13
CD351273/c
LOCUS CR139473
DEFINITION Forward strand read from insert in 3'HPRT insertion targeting and
```

```
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IWAGE:6853367"
/tissue_type="whole brain"
/dav_sage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_G10"
/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGAGCAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."
```

ORIGIN

```
Query Match 79.7%; Score 563.4; DB 6; Length 749;
Best Local Similarity 99.8%; Pred. No. 9.8e-158;
Matches 564; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 45 ATCTAGCTGATCAGGATCATATCGTCGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 104
Db 565 ATCCAGCTGATCAGGATCATATCGTCGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 506
Qy 105 TGGCGCTATCTGGGCATCGGGAGGAAGACCGTCCCTTTTCCCGAGGTTGAAGCG 164
Db 505 TGGCGCTATCTGGGCATCGGGAGGAAGACCGTCCCTTTTCCCGAGGTTGAAGCG 446
Qy 165 GCATGGAAGAGTTTCCCGAGGATGACTGCTGCTGCAATTCAGTTGAGCGAAACGACG 224
Db 445 GCATGGAAGAGTTTCCCGAGGATGACTGCTGCTGCAATTCAGTTGAGCGAAACGACG 386
Qy 225 TTTACCATGATGATTCGGGAAGGTGGCCATGACGCGCTTTAACGCTGACCTTCGTT 284
Db 385 TTTACCATGATGATTCGGGAAGGTGGCCATGACGCGCTTTAACGCTGACCTTCGTT 326
Qy 285 CAGGCCACCTGGGATACAGTTCGTCGCGCTTTTCCGGACACAGTTCCCGATGCTCAGC 344
Db 325 CAGGCCACCTGGGATACAGTTCGTCGCGCTTTTCCGGACACAGTTCCCGATGCTCAGC 266
Qy 345 CCGAAGCGCATCAGCAACCCGAAACAATACCGGGGACAGCCGGAACCTGCCGTGCGGTG 404
Db 265 CCGAAGCGCATCAGCAACCCGAAACAATACCGGGGACAGCCGGAACCTGCCGTGCGGTG 206
Qy 405 CAGATTAATGACAGCGTGGCGCTGGGATATTAATCTCAGCGAGGAGGATATCTTCGGC 464
Db 205 CAGATTAATGACAGCGTGGCGCTGGGATATTAATCTCAGCGAGGAGGATATCTTCGGC 146
Qy 465 TGGATGCGGACGAATGACATGGATACCCCGTGAGTTACCCGCGCGCGCGCTTCGTT 524
Db 145 TGGATGCGGACGAATGACATGGATACCCCGTGAGTTACCCGCGCGCGCGCTTCGTT 86
Qy 525 ATTACGCTTTTGAACCCGCTGGAGGACGGGCGAGACTCGCGGTGCAAAATGTGTTTACAGC 584
Db 85 ATTACGCTTTTGAACCCGCTGGAGGACGGGCGAGACTCGCGGTGCAAAATGTGTTTACAGC 26
Qy 585 GTGATGGAGCAGATGAGATGCTTCG 609
Db 25 GTGATGGAGCAGATGAGATGCTTCG 1
```

chromosome engineering clone MHP274b23, genomic survey sequence.

CR139473
 VERSION CR139473.1 GI:49887345
 GSS: genome survey sequence; MICR.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L., Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y., Rogers,J., and Bradley,A.
 Direct Submission
 Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. http://www.sanger.ac.uk/MICR

FEATURES
 source
 1..733
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="MHP274b23"
 /clone_lib="MHPP"

ORIGIN

Query Match 79.5%; Score 562; DB 9; Length 733;
 Best Local Similarity 99.3%; Pred. No. 2.6e-157;
 Matches 573; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 56 TCAGGATCATATCGTGGGTCCTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCT 115
 DB 733 TCAGGATCATATCGTGGGTCCTTTTCCGGCTCAGTCATCGCCCAAGCTGGCGCTATCT 674
 QY 116 GGCATCGGGAGGAGAAAGCCGCTGCTTTTCCGGAGGTTGAAGCGGCATGGAAGA 175
 DB 673 GGCATCGGGAGGAGAAAGCCGCTGCTTTTCCGGAGGTTGAAGCGGCATGGAAGA 614
 QY 176 GTTTCGGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 235
 DB 613 GTTTCGGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 554
 QY 236 GATTTCGGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 295
 DB 553 GATTTCGGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 494
 QY 296 GGATACCAAGTTCGTCGGGCTTTTCCGGACACAGTTCGGATGCTGAGCCGGAAGCGCAT 355
 DB 493 GGATACCAAGTTCGTCGGGCTTTTCCGGACACAGTTCGGATGCTGAGCCGGAAGCGCAT 434
 QY 356 CAGCAACCCGAACAATACCGCGGACAGCCGGAATGCGG -TGCGGCTGTCAGATTAAATG 414
 DB 433 CAGCAACCCGAACAATACCGCGGACAGCCGGAATGCGGCTGCTGCTGCTGCTGCTGCTGCT 374
 QY 415 ACAGCGTGGCGGCTGGGATATACGTACGAGGAGCGGATATCTGGCTGATGCGCG 474
 DB 373 ACAGCGTGGCGGCTGGGATATACGTACGAGGAGCGGATATCTGGCTGATGCGCG 314
 QY 475 AGAAATGAGATGATACCCGCTGAGTTACCCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCT 534
 DB 313 AGAAATGAGATGATACCCGCTGAGTTACCCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCT 254
 QY 535 TTGAACCCGCTGGAGGACGGGACAGCTCGCGGTGCAAAATGTTTTACAGCGTATGAGC 594
 DB 253 TTGAACCCGCTGGAGGACGGGACAGCTCGCGGTGCAAAATGTTTTACAGCGTATGAGC 194
 QY 595 AGATGAAGATGCTCGACACCGCTGAGAAACAGCGAGCT 631
 DB 193 AGATGAAGATGCTCGACACCGCTGAGAAACAGCGAGCT 157

RESULT 14
 CR167678/c
 LOCUS CR167678 633 bp DNA linear GSS 06-JUL-2004
 DEFINITION Reverse strand read from insert in 3'HPRT insertion targeting and

chromosome engineering clone MHP161k02, genomic survey sequence.

CR167678
 VERSION CR167678.1 GI:49946527
 GSS: genome survey sequence; MICR.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L., Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y., Rogers,J., and Bradley,A.
 Direct Submission
 Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. http://www.sanger.ac.uk/MICR

FEATURES
 source
 1..633
 Location/Qualifiers
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="MHP161k02"
 /clone_lib="MHPP"

ORIGIN

Query Match 78.5%; Score 555; DB 9; Length 633;
 Best Local Similarity 100.0%; Pred. No. 3.2e-155;
 Matches 555; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 TTTTTCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCATCGGGAGGAAGC 136
 DB 633 TTTTTCGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCATCGGGAGGAAGC 574
 QY 137 CCGTGCCTTTTCCCGGAGGTTGAAGCGGCATGGAAGAGTTTGCAGAGATGACTGCTG 196
 DB 573 CCGTGCCTTTTCCCGGAGGTTGAAGCGGCATGGAAGAGTTTGCAGAGATGACTGCTG 514
 QY 197 CTGCATTCAGTTGACGAAACGCAAGTTTACCATGATGATTCGGGAAGGTGGCCAT 256
 DB 513 CTGCATTCAGTTGACGAAACGCAAGTTTACCATGATGATTCGGGAAGGTGGCCAT 454
 QY 257 GCAGCGCTTTAAGGTGAAGTTCCTTCAGGCCACTCGGGATACCAAGTTTCGTCGCGCT 316
 DB 453 GCAGCGCTTTAAGGTGAAGTTCCTTCAGGCCACTCGGGATACCAAGTTTCGTCGCGCT 394
 QY 317 TTTCCGGACACAGTTCCGGATGTCAGCCGAGCGCATCAGCAACCCGGAACAATACCGG 376
 DB 393 TTTCCGGACACAGTTCCGGATGTCAGCCGAGCGCATCAGCAACCCGGAACAATACCGG 334
 QY 377 CGACAGCCGAACTGCGGTCGCGGTGTCAGATTAATGACAGCGGTGCGCGCTGGGATA 436
 DB 333 CGACAGCCGAACTGCGGTCGCGGTGTCAGATTAATGACAGCGGTGCGCGCTGGGATA 274
 QY 437 TTACGTCGAGGAGCGGATATCTCGCTGGATCGCCGAGAAATGGACATGGATACCCCG 496
 DB 273 TTACGTCGAGGAGCGGATATCTCGCTGGATCGCCGAGAAATGGACATGGATACCCCG 214
 QY 497 TGAGTTACCCGCGGCGGCTCGCTTCATTTCAGTTTGAACCCGCTGGAGGACGGGCA 556
 DB 213 TGAGTTACCCGCGGCGGCTCGCTTCATTTCAGTTTGAACCCGCTGGAGGACGGGCA 154
 QY 557 GACTCGCGGTGCAAAATGTTTACAGCGTATGAGGAGAGATGAAGATGCTCGACAGCT 616
 DB 153 GACTCGCGGTGCAAAATGTTTACAGCGTATGAGGAGAGATGAAGATGCTCGACAGCT 94
 QY 617 GCAGAACACGAGCT 631
 DB 93 GCAGAACACGAGCT 79

RESULT 15
 CR167678/c
 LOCUS CR167678 791 bp mRNA linear EST 09-JUL-2003
 DEFINITION UI-M-G10-cej-j-06-0-UI.r1 NIH_BMAP_G10 Mus musculus cDNA clone

Db	328	CAGCCACCTGGGATACCAAGTTCGTCCGGCTTTTCCGGACACAGTTCCGGATGGTCA	269
Qy	345	CCGAAGCGCATACGAAACCCGGAACAATACCGGCGACAGCCGGAACTGCCCTGCGCGTG	404
Db	268	CCGAAGCGCATACGAAACCCGGAACAATACCGGCGACAGCCGGAACTGCCCTGCGCGTG	209
Qy	405	CAGATTAAACACAGCGGTGCGGCGCTGGGATATTACGTACGCGAGGACGGGTATCCTGGC	464
Db	208	CAGATTAAACACAGCGGTGCGGCGCTGGGATATTACGTACGCGAGGACGGGTATCCTGGC	149
Qy	465	TGGATCCGCAGAAATGGACATCGATACCCCGTGAGTTACCCGGCGGGCGCGCTCGTTC	524
Db	148	TGGATCCGCAGAAATGGACATCGATACCCCGTGAGTTACCCGGCGGGCGCGCTCGTTC	89
Qy	525	ATTACAGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAAATGTTTTTAC	583
Db	88	ATTACAGTTTTTGAACCCGTGGAGGACGGGCAGACTCGCGGTGCAAAATGTTTTTAC	29
Qy	584	CGTGATGGACGATGAAGATGCTCGAC	611
Db	28	CGTGATGGACGATGAAGATGCTCGAC	1

Search completed: October 13, 2005, 06:19:24
Job time : 27488 secs

ORIGIN	Query Match	78.4%; Score 554.4; DB 6; Length 791;	
	Best Local Similarity	99.6%; Pred. No. 5e-155;	
	Matches 566; Conservative	0; Mismatches 1; Indels 1; Gaps 1;	
Qy	45	ATCTAGCTGCATCAGGATCATATCGTCGGGCTTTTTTCCGGCTCAGTCATCGCCCAAGC	104
Db	568	ATCCAGCTGCATCAGGATCATATCGTCGGGCTTTTTTCCGGCTCAGTCATCGCCCAAGC	509
Qy	105	TGGCGCTATCTGGGCATCGGGGAGGAAGCCCGTGCTTTTCCCGCGAGGTTGAAGCG	164
Db	508	TGGCGCTATCTGGGCATCGGGGAGGAAGCCCGTGCTTTTCCCGCGAGGTTGAAGCG	449
Qy	165	GCATGGAAGAAGTTTGGCGAGGATGACTGCTGCTGCATGACGTTGAGCGAAAAACGACG	224
Db	448	GCATGGAAGAAGTTTGGCGAGGATGACTGCTGCTGCATGACGTTGAGCGAAAAACGACG	389
Qy	225	TTTACCATGATGATTCGGGAAGGTGTGGCCATGACGCGCTTTTAAAGGTGAAGTCTTCGTT	284
Db	388	TTTACCATGATGATTCGGGAAGGTGTGGCCATGACGCGCTTTTAAAGGTGAAGTCTTCGTT	329
Qy	285	CAGGCCACCTGGGGATACACAGTTTCGTCGGGCTTTTTTCCGGACACAGTTCGGGATGCTCAGC	344

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 12, 2005, 14:15:25 ; Search time 4777 Seconds
(without alignments)
242.170 Million cell updates/sec

Title: US-10-826-523-40

Perfect score: 707

Sequence: 1 ggatcccatgcgtcaatttt.....aaattgacgcattggatcc 707

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	566.2	80.1	46819	3	US-09-453-702B-72 Sequence 72, Appl
2	561.4	79.4	22306	3	US-09-453-702B-251 Sequence 251, Appl
3	223	31.5	34063	3	US-09-453-702B-96 Sequence 96, Appl
4	73.8	10.4	5194	3	US-08-844-274-16 Sequence 16, Appl
5	73.8	10.4	5194	4	US-09-598-421-16 Sequence 16, Appl
6	73.4	10.4	5194	3	US-08-844-274-17 Sequence 17, Appl
7	73.4	10.4	5194	4	US-09-598-421-17 Sequence 17, Appl
8	67.8	9.6	5679	3	US-08-844-274-10 Sequence 10, Appl
9	67.8	9.6	5679	4	US-09-598-421-10 Sequence 10, Appl
10	67.8	9.6	6723	3	US-08-844-274-13 Sequence 13, Appl
11	67.8	9.6	6723	3	US-08-844-274-14 Sequence 14, Appl
12	67.8	9.6	6723	4	US-09-598-421-13 Sequence 13, Appl
13	67.8	9.6	6723	4	US-09-598-421-14 Sequence 14, Appl
14	67.8	9.6	7560	3	US-08-844-274-20 Sequence 20, Appl
15	67.8	9.6	7560	4	US-09-598-421-20 Sequence 20, Appl
16	67.8	9.6	9423	4	US-09-377-066-6 Sequence 6, Appl
17	64.4	9.1	6448	3	US-08-844-274-15 Sequence 15, Appl
18	64.4	9.1	6448	4	US-09-598-421-15 Sequence 15, Appl
19	63.8	9.0	2476	3	US-08-844-274-11 Sequence 11, Appl
20	63.8	9.0	2476	4	US-09-598-421-11 Sequence 11, Appl
21	39.8	5.6	888	4	US-09-252-991A-15689 Sequence 15689, A
22	39.8	5.6	888	4	US-09-252-991A-15722 Sequence 15722, A
23	39.8	5.6	2001	4	US-09-252-991A-15589 Sequence 15589, A
24	39.8	5.6	2475	4	US-09-252-991A-15758 Sequence 15758, A
25	39.2	5.5	1734	6	Patent No. 5352575
26	39.2	5.5	1734	6	Patent No. 5352575-8
27	35.4	5.0	209210	4	US-09-949-016-15094 Sequence 15094, A

ALIGNMENTS

RESULT 1

US-09-453-702B-72

; Sequence 72, Application US/09453702B

; Patent No. 6365723

; GENERAL INFORMATION:

; APPLICANT: Blattner, Frederick R.

; Burland, Nicole T.

; Plunkett, Guy

; Welch, Rod

; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157

; NUMBER OF SEQUENCES: 265

; CORRESPONDENCE ADDRESS:

; ADDRESSES: Quarles & Brady

; STREET: 1 South Pinckney Street

; CITY: Madison

; STATE: WI

; COUNTRY: US

; ZIP: 53701-2113

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB storage

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Word Perfect 8.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/453,702B

; FILING DATE: 03-Dec-1999

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/110,955

; FILING DATE: 04-DEC-1998

; ATTORNEY/AGENT INFORMATION:

; NAME: Seay, Nicholas J.

; REGISTRATION NUMBER: 27386

; REFERENCE/DOCKET NUMBER: 960296.95017

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (608) 251-5000

; TELEFAX: (608) 251-9166

; INFORMATION FOR SEQ ID NO: 72:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 46819

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; SEQUENCE DESCRIPTION: SEQ ID NO: 72:

US-09-453-702B-72

Query Match 80.1%; Score 566.2; DB 3; Length 46819;

Best Local Similarity 97.8%; Pred. No. 6.6e-176;

Sequence 3154, Ap
Sequence 2816, Ap
Sequence 2, Appl
Sequence 1, Appl
Sequence 7848, Ap
Sequence 1011, Ap
Sequence 1011, Ap
Sequence 1011, Ap
Sequence 1011, Ap
Sequence 785, App
Sequence 1230, Ap
Sequence 573, App
Sequence 1868, Ap
Sequence 17150, A
Sequence 2, Appl
Sequence 2, Appl
Sequence 2, Appl

Matches 574; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 104
DB 14795 ATCCAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 14854
QY 105 TGGCGCTATCTGGGCATCGGGAGGAAGAAGCCCGTGCCTTTTTCGGCGAGGTTGAAGCG 164
DB 14855 TGGCGCTATCTGGGCATCGGGAGGAAGAAGCCCGTGCCTTTTTCGGCGAGGTTGAAGCG 14914
QY 165 GCATGGAAGAGTTTGGCAGGATGACTGCTGCTGCATTTAGCGGAAACGACG 224
DB 14915 GCATGGAAGAGTTTGGCAGGATGACTGCTGCTGCATTTAGCGGAAACGACG 14974
QY 225 TTACCATGATGATTGGGAGGTGTGGCATGACGCTTTTAAACGGTGAACGTTCGTT 284
DB 14975 TTACCATGATGATTGGGAGGTGTGGCATGACGCTTTTAAACGGTGAACGTTCGTT 15034
QY 285 CAGGCCACCTGGGATACCGTCCCTCGGACTGTTTCGGGACACAGTTCGGGATGTCAGC 344
DB 15035 CAGGCCACCTGGGATACCGTCCCTCGGACTGTTTCGGGACACAGTTCGGGATGTCAGC 15094
QY 345 CCGAAGCCCATCAGCAACCCGAAACAATACCGGCGACGACCGGAACTGCCGTGCGGTG 404
DB 15095 CCGAAGCCCATCAGCAACCCGAAACAATACCGGCGACGACCGGAACTGCCGTGCGGTG 15154
QY 405 CAGATTATGACAGCGGTGCGGCGCTGGGATATTACGTGACGAGGAGCGGTATCCTGGC 464
DB 15155 CAGATTATGACAGCGGTGCGGCGCTGGGATATTACGTGACGAGGAGCGGTATCCTGGC 15214
QY 465 TGGATGCCGCGAAGATGACATGATACCCCGTGGATTACCGGCGGCGGCGCTCGTTC 524
DB 15215 TGGATGCCGCGAAGATGACATGATACCCCGTGGATTACCGGCGGCGGCGCTCGTTC 15274
QY 525 ATTACGCTTTTGAACCCGTGGAGGACCGGCGAGCTCGCGGTGCAATGTGTTTACAGC 584
DB 15275 ATTACGCTTTTGAACCCGTGGAGGACCGGCGAGCTCGCGGTGCAATGTGTTTACAGC 15334
QY 585 GTGATGGAGCAGATGCAAGATGCTGCACACGCTGCAGAACACGCGACT 631
DB 15335 GTGATGGAGCAGATGCAAGATGCTGCACACGCTGCAGAACACGCGACT 15381

RESULT 2

US-09-453-702B-251/c
; Sequence 251, Application US/09453702B
; Patent No. 6365723

GENERAL INFORMATION:

APPLICANT: Blattner, Frederick R.

Burland, Valerie

Perna, Nicole T.

Plunkett, Guy

Welch, Rod

TITLE OF INVENTION: No. 6365723el Sequences of E. coli 0157

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESSEE: Quarles & Brady

STREET: 1 South Pinckney Street

CITY: Madison

STATE: WI

COUNTRY: US

ZIP: 53701-2113

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 8.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/453,702B

FILING DATE: 03-Dec-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/110,955

; FILING DATE: 04-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 251:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22306
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 251:
US-09-453-702B-251

Query Match 79.4%; Score 561.4; DB 3; Length 22306;
Best Local Similarity 97.3%; Pred. No. 1.7e-174;
Matches 571; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 45 ATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 104
DB 17193 ATCCAGTTGCATCAGGATCATATCGTCGGGTCTTTTTCGGGCTCAGTCATCGCCCAAGC 17134
QY 105 TGGCGCTATCTGGGCATCGGGAGGAAGAAGCCCGTGCCTTTTTCGGCGAGGTTGAAGCG 164
DB 17133 TGGCGCTATCTGGGCATCGGGAGGAAGAAGCCCGTGCCTTTTTCGGCGAGGTTGAAGCG 17074
QY 165 GCATGGAAGAGTTTGGCAGGATGACTGCTGCTGCATTTAGCGGAAACGCAAGC 224
DB 17073 GCATGGAAGAGTTTGGCAGGATGACTGCTGCTGCATTTAGCGGAAACGCAAGC 17014
QY 225 TTACCATGATGATTGGGAGGTGTGGCATGACGCTTTTAAACGGTGAACGTTCGTT 284
DB 17013 TTACCATGATGATTGGGAGGTGTGGCATGACGCTTTTAAACGGTGAACGTTCGTT 16954
QY 285 CAGGCCACCTGGGATACCGTCCCTCGGACTGTTTCGGGACACAGTTCGGGATGTCAGC 344
DB 16953 CAGGCCACCTGGGATACCGTCCCTCGGACTGTTTCGGGACACAGTTCGGGATGTCAGC 16894
QY 345 CCGAAGCCCATCAGCAACCCGAAACAATACCGGCGACGACCGGAACTGCCGTGCGGTG 404
DB 16893 CCGAAGCCCATCAGCAACCCGAAACAATACCGGCGACGACCGGAACTGCCGTGCGGTG 16834
QY 405 CAGATTATGACAGCGGTGCGGCGCTGGGATATTACGTGACGAGGAGCGGTATCCTGGC 464
DB 16833 CAGATTATGACAGCGGTGCGGCGCTGGGATATTACGTGACGAGGAGCGGTATCCTGGC 16774
QY 465 TGGATGCCGCGAAGATGACATGATACCCCGTGGATTACCGGCGGCGGCGCTCGTTC 524
DB 16773 TGGATGCCGCGAAGATGACATGATACCCCGTGGATTACCGGCGGCGGCGCTCGTTC 16714
QY 525 ATTACGCTTTTGAACCCGTGGAGGACCGGCGAGCTCGCGGTGCAATGTGTTTACAGC 584
DB 16713 ATTACGCTTTTGAACCCGTGGAGGACCGGCGAGCTCGCGGTGCAATGTGTTTACAGC 16654
QY 585 GTGATGGAGCAGATGCAAGATGCTGCACACGCTGCAGAACACGCGACT 631
DB 16653 GTGATGGAGCAGATGCAAGATGCTGCACACGCTGCAGAACACGCGACT 16607

RESULT 3

US-09-453-702B-96/c
; Sequence 96, Application US/09453702B
; Patent No. 6365723

GENERAL INFORMATION:

APPLICANT: Blattner, Frederick R.

Burland, Valerie

Perna, Nicole T.

Plunkett, Guy

Welch, Rod

; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
 ; NUMBER OF SEQUENCES: 265
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Quarles & Brady
 ; STREET: 1 South Pinckney Street
 ; CITY: Madison
 ; STATE: WI
 ; COUNTRY: US
 ; ZIP: 53701-2113
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Word Perfect 8.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/453,702B
 ; FILING DATE: 03-Dec-1999
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/110,955
 ; FILING DATE: 04-DEC-1998
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Seay, Nicholas J.
 ; REGISTRATION NUMBER: 27386
 ; REFERENCE/DOCKET NUMBER: 960296.95017
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (608) 251-5000
 ; TELEFAX: (608) 251-9166
 ; INFORMATION FOR SEQ ID NO: 96:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 34063
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 96:
 US-09-453-702B-96

Query Match 31.5%; Score 223; DB 3; Length 34063;
 Best Local Similarity 61.1%; Pred. No. 2.6e-62;
 Matches 361; Conservative 0; Mismatches 230; Indels 0; Gaps 0;
 QY 51 CTGCATCAGGATCATATCGTGGGTCCTTTTTCGGGTCAGTCATCGCCCAAGCTGGCGC 110
 DB 14695 CTGCATAAGGATCATTGTCGGGCATATGTTCTTATCAGTCACGTCGGAAGCTGGCGC 14636
 QY 111 TATCTGGGATCGGGAGGAGAAAGCCCGTCCTTTTCCCGCGAGGTTGAAGCGGCATGG 170
 DB 14635 TGGCTGGGGATGCGGGAGACCGCAGCAAAAGCTTTTGTGATGAGGTGGAGCGGCGCTGG 14576
 QY 171 AAAGAGTTTCCGAGGATGACTGCTGCTGATTGACGTTGAGCGGAAACGACGTTTACC 230
 DB 14575 TCGGAATACGCCGAAGGAGTGTCTGGGAGATGACGTGGAAGAAACGACGTTTACG 14516
 QY 231 ATGATGATTCGGGAAGGTGGCCATGCGCCCTTTTAAACGGTGAACCTGTTCTGTTTCAGGCC 290
 DB 14515 GAATTTATCCGTGAAGGTGGGGGTTTCATCGGTTTAAACGGGAAATCTTTGTGACGCG 14456
 QY 291 ACCTGGGATACCAAGTTCGTGCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAG 350
 DB 14455 GTCTGGGATACGGAACCAAGCAGTATTTCGTACGGGTTTTTAAAGCGGTGAGTCCGAAA 14396
 QY 351 CGCATCAGCAACCGAACAATACCGGACGAGCGGAACTGCGTCCGGGTGGTGCAGATT 410
 DB 14395 CGGGTGGACACCGCAGGACACGATGATGGGAAACCGTTTCTGCGGGCGGGGTGGAGGTC 14336
 QY 411 AATCAGACGCGTGGCGCTGGGATATATCAGTCAGGAGGACGGGTATCTCTGGCTGGATG 470
 DB 14335 GATCGATATGCGCGTGGTTCGATACCATATCTGTGAGGATGATTTTCTCGCTCCGGG 14276
 QY 471 CCGCAGAAATGGAATGATGATACCCCGTGAAGTTACCCCGCGGCGCGCTGTTCAATCAC 530
 DB 14275 AGTGGACGATGGGAACGATCCCGCGTGAACTTCCACCCGCGGCGTCCGGCCATGCTGCAT 14216

QY 531 GTTTTGAACCCGTGGAGACCGGCGAGACTCGGGTGCAATGTGTTTACACGCTGATG 590
 DB 14215 ATTTTCGAGCCGCTGGAGGACCGGCGAGACCCGTGGGGCCCAACCCAGTTTACACGCTGATG 14156
 QY 591 GAGCAGATGAAGATGCTCGACACGCTGCGAGAAACGACGCTAGATTAAACC 641
 DB 14155 GAACGCTGAAGATGCTCGATTCCCTGCGAGGCAACACAGCTTCAGTCGGGC 14105
 RESULT 4
 US-08-844-274-16/c
 ; Sequence 16, Application US/08844274B
 ; Patent No. 6218185
 ; GENERAL INFORMATION:
 ; APPLICANT: Fraser Jr., Malcom J.
 ; APPLICANT: Shirk, Paul D.
 ; APPLICANT: Elick, Teri A.
 ; APPLICANT: Perera, Omathage
 ; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
 ; TITLE OF INVENTION: for Insects
 ; FILE REFERENCE: 0148.96
 ; CURRENT APPLICATION NUMBER: US/08/844,274B
 ; EARLIER FILING DATE: 1997-04-18
 ; EARLIER APPLICATION NUMBER: 60/016,234
 ; EARLIER FILING DATE: 1996-04-19
 ; NUMBER OF SEQ ID NOS: 22
 ; SOFTWARE: PatentIn Ver. 2.0 - beta
 ; SEQ ID NO 16
 ; LENGTH: 5194
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: plasmid
 ; OTHER INFORMATION: IFF2B/Xpuc18.1
 US-08-844-274-16

Query Match 10.4%; Score 73.8; DB 3; Length 5194;
 Best Local Similarity 83.2%; Pred. No. 2.3e-13;
 Matches 84; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
 QY 605 GCTCGACACGCTCGACAAACGACGAGTAACTTAAACCTAGAAAGATAATCATATTGTGAC 664
 DB 2955 GCTCGGATCCGGGGATCGGATCTCTAGATTAAACCTAGAAAGATAATCATATTGTGAC 2896
 QY 665 GTACGTTAAAGATAAATCATGCGTAAATTCAGCATGGGAT 705
 DB 2895 GTACGTTAAAGATAAATCATGCGTAAATTCAGCATGGT 2855

RESULT 5
 US-09-598-421-16/c
 ; Sequence 16, Application US/09598421
 ; Patent No. 6551825
 ; GENERAL INFORMATION:
 ; APPLICANT: Fraser Jr., Malcom J.
 ; APPLICANT: Shirk, Paul D.
 ; APPLICANT: Elick, Teri A.
 ; APPLICANT: Perera, Omathage
 ; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
 ; TITLE OF INVENTION: for Insects
 ; FILE REFERENCE: 0148.96
 ; CURRENT APPLICATION NUMBER: US/09/598,421
 ; CURRENT FILING DATE: 2000-06-19
 ; PRIOR APPLICATION NUMBER: 60/016,234
 ; PRIOR FILING DATE: 1996-04-19
 ; NUMBER OF SEQ ID NOS: 22
 ; SOFTWARE: PatentIn Ver. 2.0 - beta
 ; SEQ ID NO 16
 ; LENGTH: 5194
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:

```
; OTHER INFORMATION: Description of Artificial Sequence:plasmid
; OTHER INFORMATION: IFF2B/Xpuc18.1
US-09-598-421-16

Query Match          10.4%; Score 73.8; DB 4; Length 5194;
Best Local Similarity 83.2%; Pred. No. 2.3e-13;
Matches 84; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 605 GCTCGACACGGCTGCAGACACGACGCTAGATTAAACCCCTAGAAAGATAATCATATTGTGCAC 664
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2955 GCTCGTACCCGGGATCGGATCCCTCTAGATTAAACCCCTAGAAAGATAATCATATTGTGCAC 2896

Qy 665 GTACGTTAAAGATAATCATGCGTAAAAATTGACGCATGGGAT 705
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2895 GTACGTTAAAGATAATCATGCGTAAAAATTGACGCATGTGTT 2855

RESULT 6
US-08-844-274-17/c
; Sequence 17, Application US/08844274B
; Patent No. 6218185
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omathhage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic transformation System
; TITLE OF INVENTION: for Insects
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/08/844,274B
; CURRENT FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/016,234
; EARLIER FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 17
; LENGTH: 5194
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:plasmid
; OTHER INFORMATION: IFF2B/XsupF4H
US-08-844-274-17

Query Match          10.4%; Score 73.4; DB 3; Length 5194;
Best Local Similarity 87.9%; Pred. No. 3.1e-13;
Matches 80; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 615 CTGCAGAACACGCGCTAGATTAAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAA 674
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4773 CTCTAGAGGGATCCTCTAGATTAAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAA 4714

Qy 675 GATAATCATGCGTAAAAATTGACGCATGGGAT 705
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4713 GATAATCATGCGTAAAAATTGACGCATGTGTT 4683

RESULT 7
US-09-598-421-17/c
; Sequence 17, Application US/09598421
; Patent No. 6551825
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omathhage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic transformation System
; TITLE OF INVENTION: for Insects
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/09/598,421
; CURRENT FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 60/016,234
; PRIOR FILING DATE: 1996-04-19
```

```
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 17
; LENGTH: 5194
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:plasmid
; OTHER INFORMATION: IFF2B/XsupF4H
US-09-598-421-17

Query Match          10.4%; Score 73.4; DB 4; Length 5194;
Best Local Similarity 87.9%; Pred. No. 3.1e-13;
Matches 80; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 615 CTGCAGAACACGCGCTAGATTAAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAA 674
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4773 CTCTAGAGGGATCCTCTAGATTAAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAA 4714

Qy 675 GATAATCATGCGTAAAAATTGACGCATGGGAT 705
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4713 GATAATCATGCGTAAAAATTGACGCATGTGTT 4683

RESULT 8
US-08-844-274-10
; Sequence 10, Application US/08844274B
; Patent No. 6218185
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omathhage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic transformation System
; TITLE OF INVENTION: for Insects
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/08/844,274B
; CURRENT FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/016,234
; EARLIER FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 10
; LENGTH: 5679
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:p3E1.2 delta
; OTHER INFORMATION: TRL
US-08-844-274-10

Query Match          9.6%; Score 67.8; DB 3; Length 5679;
Best Local Similarity 97.2%; Pred. No. 2.3e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 635 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATGCGTAAATTTG 694
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1080 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATGCGTAAATTTG 1139

Qy 695 AGCGATGGGAT 705
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1140 AGCGATGTGTT 1150

RESULT 9
US-09-598-421-10
; Sequence 10, Application US/09598421
; Patent No. 6551825
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omathhage
```

```
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; TITLE OF INVENTION: for Insects
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/09/598,421
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 60/016,234
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 10
; LENGTH: 5679
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:p3el.2 delta
; OTHER INFORMATION: TRL
US-09-598-421-10

Query Match          9.6%; Score 67.8; DB 4; Length 5679;
Best Local Similarity 97.2%; Pred. No. 2.5e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 635 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 694
Db 1080 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 1139

QY 695 ACGCATGGGAT 705
Db 1140 ACGCATGTGTT 1150

RESULT 10
US-08-844-274-13
; Sequence 13, Application US/08844274B
; Patent No. 6218185
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omaththage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/08/844,274B
; PRIOR FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/016,234
; EARLIER FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 13
; LENGTH: 6723
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Clone p3el.2H/S
US-08-844-274-13

Query Match          9.6%; Score 67.8; DB 3; Length 6723;
Best Local Similarity 97.2%; Pred. No. 2.5e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 635 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 694
Db 1080 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 1139

QY 695 ACGCATGGGAT 705
Db 1140 ACGCATGTGTT 1150

RESULT 11
US-08-844-274-14/c
; Sequence 14, Application US/08844274B
```

```
; Patent No. 6218185
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omaththage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/08/844,274B
; PRIOR FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/016,234
; EARLIER FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 14
; LENGTH: 6723
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:piggyBac/opd
US-08-844-274-14

Query Match          9.6%; Score 67.8; DB 3; Length 6723;
Best Local Similarity 97.2%; Pred. No. 2.5e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 635 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 694
Db 5644 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 5585

QY 695 ACGCATGGGAT 705
Db 5584 ACGCATGTGTT 5574

RESULT 12
US-09-598-421-13
; Sequence 13, Application US/09598421
; Patent No. 6551825
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Elick, Teri A.
; APPLICANT: Perera, Omaththage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic Transformation System
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/09/598,421
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 60/016,234
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 13
; LENGTH: 6723
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Clone p3el.2H/S
US-09-598-421-13

Query Match          9.6%; Score 67.8; DB 4; Length 6723;
Best Local Similarity 97.2%; Pred. No. 2.5e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 635 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 694
Db 1080 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 1139

QY 695 ACGCATGGGAT 705
Db 1140 ACGCATGTGTT 1150
```

```
RESULT 13
US-09-598-421-14/c
; Sequence 14, Application US/09598421
; Patent No. 6551825
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Ellick, Teri A.
; APPLICANT: Perera, Omathchage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic transformation System
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/09/598,421
; CURRENT FILING DATE: 2000-06-19
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 14
; LENGTH: 6723
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:piggyBac/opd
US-09-598-421-14

Query Match          9.6%; Score 67.8; DB 4; Length 6723;
Best Local Similarity 97.2%; Pred. No. 2.5e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 635 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 694
Db 5644 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 5585

Qy 695 ACGCATGGGAT 705
Db 5584 ACGCATGTGTT 5574

RESULT 14
US-08-844-274-20/c
; Sequence 20, Application US/08844274B
; Patent No. 6218185
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Ellick, Teri A.
; APPLICANT: Perera, Omathchage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic transformation System
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/08/844,274B
; CURRENT FILING DATE: 1997-04-18
; EARLIER APPLICATION NUMBER: 60/016,234
; EARLIER FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 20
; LENGTH: 7560
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:p3E1.2hs/opd
US-08-844-274-20

Query Match          9.6%; Score 67.8; DB 3; Length 7560;
Best Local Similarity 97.2%; Pred. No. 2.7e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 635 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 694
Db 5644 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 694

Qy 695 ACGCATGGGAT 705
Db 5584 ACGCATGTGTT 5574

RESULT 15
US-09-598-421-20/c
; Sequence 20, Application US/09598421
; Patent No. 6551825
; GENERAL INFORMATION:
; APPLICANT: Fraser Jr., Malcom J.
; APPLICANT: Shirk, Paul D.
; APPLICANT: Ellick, Teri A.
; APPLICANT: Perera, Omathchage
; TITLE OF INVENTION: PiggyBac Transposon-Based Genetic transformation System
; FILE REFERENCE: 0148.96
; CURRENT APPLICATION NUMBER: US/09/598,421
; CURRENT FILING DATE: 2000-06-19
; PRIOR FILING DATE: 1996-04-19
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0 - beta
; SEQ ID NO 20
; LENGTH: 7560
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:p3E1.2hs/opd
US-09-598-421-20

Query Match          9.6%; Score 67.8; DB 4; Length 7560;
Best Local Similarity 97.2%; Pred. No. 2.7e-11;
Matches 69; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 635 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 694
Db 5644 TTAACCCCTAGAAAGATAATCATATTGTGACGTACGTTAAAGATAATCATCGTAAATTTG 694

Qy 695 ACGCATGGGAT 705
Db 5584 ACGCATGTGTT 5574

Search completed: October 13, 2005, 07:39:19
Job time : 4784 secs
```